MAY 1 1 1982

1051-H11.

Y 3. N88: 25/2434

NUREG/CR-2434
LA-9116-MS

COMPLETED ORIGINAL

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36.

FRAC (Failure Rate Analysis Code)
A Computer Program for Analysis of Variance of Failure Rates
An Application Users' Guide

Los Alamos Los Alamos National Laboratory
Los Alamos, New Mexico 87545

An Affirmative Action/Equal Opportunity Employer

Typed by Jan Stelzer, S-DO

NOTICE

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, or any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for any third party's use, or the results of such use, of any information, apparatus, product or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights.

NUREG/CR-2434
LA-9116-MS

AN, XA, and RG

FRAC (Failure Rate Analysis Code): A Computer Program for Analysis of Variance of Failure Rates

An Application User's Guide

H. F. Martz
R. J. Beckman
C. R. McInteer

Manuscript submitted: February 1982
Date published: March 1982

Prepared for
Division of Risk Analysis
Office of Nuclear Regulatory Research
US Nuclear Regulatory Commission
Washington, DC 20555

NRC FIN No. A7725

Los Alamos National Laboratory
Los Alamos, New Mexico 87545

FRAC (FAILURE RATE ANALYSIS CODE): A COMPUTER PROGRAM FOR ANALYSIS OF VARIANCE OF FAILURE RATES-- AN APPLICATION USER'S GUIDE

by

H. F. Martz, R. J. Beckman, and C. R. McInteer

ABSTRACT

Probabilistic risk assessments (PRAs) require estimates of the failure rates of various components whose failure modes appear in the event and fault trees used to quantify accident sequences. Several reliability data bases have been designed for use in providing the necessary reliability data to be used in constructing these estimates. In the nuclear industry, the Nuclear Plant Reliability Data System (NPRDS) and the In-Plant Reliability Data System (IRPDS), among others, were designed for this purpose.

An important characteristic of such data bases is the selection and identification of numerous factors used to classify each component that is reported and the subsequent failures of each component. However, the presence of such factors often complicates the analysis of reliability data in the sense that it is inappropriate to group (that is, pool) data for those combinations of factors that yield significantly different failure rate values. These types of data can be analyzed by analysis of variance. Analysis of variance is a statistical data analysis methodology for use in addressing such questions as: How do the factors affect the failure rate? What are the estimated effects due to these factors? Which factor combinations yield the largest failure rate estimates? Are there factor interactions that significantly affect the failure rate? FRAC (Failure Rate Analysis Code) is a computer code that performs an analysis of variance of failure rates and provides information for answering the above questions. In addition, FRAC provides failure rate estimates.

I. INTRODUCTION

The purpose of this report is to provide a description of a computer code called FRAC (Failure Rate Analysis Code) that can be used to simultaneously investigate the combined effects of various factors on component failure rates.

FRAC is useful in mainly two situations; namely, (1) when reliability failure data exist for a class of components from a structured data collection effort (a data base) that must be analyzed, and (2) when component failure data exist for various (non-identical) use conditions that are to be "pooled" (or "grouped") as a means of increasing the quantity of failure data while recognizing that the failure rate values may perhaps be different depending on the effect of the use conditions on the failure rate.

FRAC recognizes and incorporates numerous features that make it an extremely versatile data analysis code. FRAC analyzes incomplete data sets. By "incomplete," we mean that there are numerous factor combinations for which no failures have been observed or for which no components are present in the population. Thus, the data available for such an analysis are usually unbalanced in a statistical sense. FRAC analyzes both fixed and random factors. FRAC also analyzes both factorial, as well as nested (hierarchical), factors. Point estimates, as well as both confidence and tolerance interval failure rate estimates, are produced. In addition, FRAC provides failure rate estimates for components in which the values of all the factors are not all known. This often occurs in performing PRAs for proposed plants in which precise component design specifications have not yet been finalized.

FRAC utilizes a log-linear model in which all random factor effects, as well as the random error, are assumed to have log-normal distributions. Maximum likelihood estimation methods are used to estimate the factor effects. Confidence and tolerance interval estimates are based on error propagation techniques. Both FRAC input and output are described in detail, and numerous examples illustrate the practical implementation of FRAC.

The reader is assumed to have some familiarity with the statistical topic known as analysis of variance. A brief introduction to the use of an analysis of variance approach in estimating failure rates is given by Martz and Beckman.[1] Additional and more thorough background information may be found in many textbooks, such as Bowker and Lieberman,[2] Hicks,[3] and Scheffe.[4]

The topic "analysis of variance" falls within the general statistical heading known as "experimental design" or the "design of experiments."

In order to quantify the risk in operating potentially hazardous systems such as nuclear power plants, fault trees and event trees are frequently used to synthesize the complex pattern of events accompanying various postulated accident scenarios. Most of the bottommost events in such analyses, the so-called basic events, concern the failure characteristics of components of safety systems designed to mitigate the occurrence of such accidents. It is commonly required to estimate the failure rate per unit time associated with various time-dependent modes of component failure, as well as the failure rate per demand (that is, the probability of failure per demand) for demand-dependent component failure modes. For example, the reliability of a core auxiliary cooling system (CACS) of a nuclear power plant depends on the time-dependent rate of failure while on standby of such components as the auxiliary loop isolation valve used to isolate this loop of the CACS. As another example, the reliability of the emergency on-site power supply of a nuclear plant depends on the failure of the diesel-powered generators to start on demand.

In the discussion that follows, it is assumed that it is required to estimate a component failure rate for a certain well-defined type of failure, such as a catastrophic failure, as well as a well-defined mode of failure (or perhaps several or all modes combined) of a component of interest.

It is further assumed that all failure rates are constant during the time period for which the failure rate is to be estimated and for which the failure data have been collected. Thus, for any given component, the number of failures R in time T follows a Poisson distribution with parameter λT. That is,

$$\Pr(R = r) = \frac{(\lambda T)^r}{r!} \exp(-\lambda T), \; r = 0, 1, 2, \ldots, \; \lambda, T > 0 \quad , \tag{1}$$

where λ is the underlying failure rate of interest to be estimated and analyzed. However, the procedure to be presented can be used if λ is <u>not</u> constant by partitioning the observation period T into time-ordered subperiods $T_1, T_2, \ldots, T_p$ (where $T = \Sigma T_i$) with corresponding (constant) failure rates $\lambda_1, \lambda_2, \ldots, \lambda_p$, respectively. The procedure can then be used in conjunction with the observed number of failures r_i in subperiod

T_i to estimate the sequence of λ_i values that are then examined for any time-dependent trends, etc.

In the case of demand-dependent failure modes, the data consist of the observed number of failures r in n operating and/or test demands. If the probability of failure per demand λ (the failure rate per demand) is small, the exact binomial distribution of R is approximated by the Poisson distribution given in Eq. (1) in which T is replaced by n.

It is well known that component failure rates often depend on specific use conditions. For example, the same pump may possess a different failure rate due to such things as plant-to-plant differences, system differences, environmental effects, maintenance and operating effects, etc. The purpose of the procedure to be described is to account for and to estimate the simultaneous effects of such factors on component failure rates.

It is customary in many reliability data bases to identify the particular values for a selected set of such factors and to associate these factor values with observed failures of the device. For example, the In-Plant Reliability Data System (IPRDS) identifies, for each component in the population, such factors as plant Nuclear Steam Supply System (NSSS) vendor, system use, component vendor, a hierarchical classification breakdown of component design, component operating and size qualifiers, and mode of operation. Such data are supplied by the plant availability data and population data records. For instance, consider the component category "pumps" as reported in the data base. For each pump in the plants surveyed, such factors as plant ID, NSSS vendor ID (4 choices), system ID (45 choices), component vendor ID, and hierarchical pump design breakdown are reported (see Table I). Finally, the size and operating qualifiers, namely pump head size (4 categories), horsepower (6 categories), and operating temperature (3 categories) are also reported for each pump. In this example it may be desirable to determine which of these 10 main classification factors (and possible interactions of these factors) significantly affect the failure rate of pumps based on the data-base failure data reported for a set of such pumps. For a discussion of interaction factors, see Ref. 1. It is also usually desirable to estimate the failure rate of such pumps for various desired combinations (called "levels") of such factors corresponding to either known use conditions or hypothetical (planned) use conditions (for perhaps nonexistent but proposed plants). It is also often the case in practice that a pump failure rate estimate must be provided

TABLE I

PUMP FACTORS

Primary Factor	Secondary Factor	Tertiary Factor
Positive Displacement	Reciprocating	Axial Piston (Swash Plate) Piston/Crank Shaft
	Rotary	Vane Lobed Impeller (Boots Type) Gear Type Wankle
	Radial Plunger	
Centrifugal	Turbine	
	Volute	Single Suction Double Suction

in which not all of the factors, such as those listed above, are known. For example, it may be desired to estimate the failure rate of a 100-hp positive displacement pump. In this case, only the values of two factors out of the 10 possible main factors are specified and the values of the remaining eight main factors are unknown. The procedure to be described can also be used to estimate failure rates in such cases that are characterized by incomplete (either unknown or unspecified) population and/or operating conditions.

If an analysis is to be performed for determining any time-dependent trends in the failure rate λ, then "time" would be considered as yet another "factor" and subdivided (partitioned) into subperiods T_1, T_2, ... , T_p. The subperiods would denote the p "levels" of the "time factor." The combined failure data for the total observed time T would then be correspondingly partitioned, and the effect of time could be analyzed as though it were any other factor, such as those listed above. However, in the remainder of this report, it will be assumed that the failure rate is constant for the total accumulated observation/test time T and thus, time will not be considered as a factor in the analysis.

In the past it has been customary to examine the effect of such factors one at a time. This is an inefficient method of analysis that does not lend itself to convenient investigation of interaction effects and can also be

incorrect. For example, suppose we are investigating the effect of size of pump on the failure rate. If data are pooled in the data base into each of several size categories, the apparent significant effect of size may really be due to other factors, such as pump vendor, which are either unknown or unreported ("summed out"). Such "confounding" can result in often confusing, and sometimes contradictory, reports regarding the significant factors that affect pump failure rate. This situation can usually be avoided by simultaneous consideration of either several or all of the main factors in a single combined analysis. The statistical name "analysis of variance" is used to denote the combined analysis of the effects of several factors on a response variable (in this case the estimated failure rate). Thus, the procedure to be described and the associated computer code are known as FRAC (Failure Rate Analysis Code).

Such an analysis of variance procedure must be compatible with the type and quantity of reliability data available for analysis in reliability data bases such as NPRDS, LERS, IPRDS, etc. There are several characteristic features of most of these data bases:

(1) Population and operating environmental factors (such as those exemplified above) are usually each partitioned into categories (the possible levels of each factor) for both continuous factors (such as size or horsepower) and discrete factors (such as component vendor ID). This is usually done as a convenient aid in coding the data.

(2) Most of the data reported will fall into only a small fraction (perhaps 10% or even less) of all the total (multiplicative) combinations of levels of all the factors. Many cells (possible combinations of all factors) will be technically infeasible and no data will ever be reported in such cells.

(3) Many cells exist for which no failures have been observed in a stated operating time. This is due to the intrinsic high reliability of the components of such plants.

(4) Some of the factors are best interpreted as random factors in the sense to be discussed in Sec. II.

The procedure to be described recognizes such characteristics of the data-base information used to perform the failure rate analysis and is fully compatible with these (and other) features of the data.

The output of FRAC consists of several kinds and types of component failure rate estimates depending on the options selected by the analyst. The most common estimates include:

(1) the estimated fixed factor effects on the failure rate, as well as their variances and covariances;

(2) estimates of the variances of all of the random effects, including the random error term, as well as the variances and covariances of these variance estimates;

(3) point estimates of the median failure rate for each of the factor combinations (the cells) present in the input data;

(4) confidence interval estimates of the failure rate for each of the cells present in the input data;

(5) tolerance interval estimates on either the population of failure rates (in the case where there are random effects present) or the population of failure rate estimates corresponding to each of the cells present in the input data for a user-specified time interval; and

(6) point and tolerance interval estimates for any user-specified set of cells and average estimates for this set.

A complete discussion of each of these estimates and their construction is given in Sec. III.B based on the model to be presented and discussed in Sec. II. Martz and Beckman[1] provide a general expository introduction to the use of an analysis of variance approach to failure rate estimation and discuss the distinctions between fixed and random factors, as well as confidence and tolerance intervals. The report by Martz and Beckman[1] is considered to be a companion introductory tutorial document to this report and should be obtained by those readers desiring additional background on the statistical and, in particular, the analysis of variance basis of failure rate estimation.

Section II discusses the model and assumptions upon which FRAC is based. Section III is an application user's guide to the use of FRAC, including a description of the input for setting up and using FRAC (Sec. III.A), and a description and discussion of the output (Sec. III.B). A programmer's guide to FRAC is also available (see Ref. 5).

II. MODEL AND ASSUMPTIONS

It is assumed that the underlying failure rate λ is a function of K factors $F_1, F_2, \ldots, F_k$. The specific K factors may vary depending on the classification scheme used in constructing the corresponding data base.

The levels of factor F_j are denoted as 1, 2, ... , m_j, j = 1, 2, ... , K. For example, if F_1 denotes "plant ID," and if there are three plants, X, Y, and Z, represented in the data base, then F_1 has three levels indexed as 1, 2, and 3. Thus, if a particular component is from plant Y, say, then the level of F_1 for this component is indicated as level 2, and the integer "2" would be identified as the "value" (that is, the specific level) of F_1 for this component.

We further denote the specific levels of the K factors associated with the s^{th} cell (that is, the s^{th} Poisson data set) as the set of pairs of indices $S = \{(1,s(1)),(2,s(2)), \ldots , (K,s(K))\}$ where s(j) is an integer from the set $\{1,2, \ldots , m_j\}$. For example, if K = 3, then $S = \{(1,2),(2,3),(3,1)\}$ indicates that F_1 is at level 2, F_2 is at level 3, and F_3 is at level 1 corresponding to that cell. Further examples are given in Sec. III.B.

The particular model assumed in FRAC for the failure rate λ_s associated with the s^{th} cell for which data are available is

$$\lambda_s = \lambda_g \prod_{j=1}^{K} A(j,s(j)) \quad , \tag{2}$$

where A(j,s(j)) represents the multiplicative effect of factor F_j at the $s(j)^{th}$ level on a generic failure rate λ_g. The generic failure rate λ_g can be thought of as the geometric average of the failure rate values over all the use conditions represented by the data used in the analysis. Thus, λ in Eq. (1) is assumed to have the multiplicative form given in Eq. (2) for the s^{th} cell.

For the s^{th} cell considered in the analysis, the underlying failure rate λ_s may be estimated by means of the usual well-known estimator given by

$$\hat{\lambda}_s = \frac{r_s}{T_s} \quad . \tag{3}$$

It is further assumed that $\hat{\lambda}_s$ is related to λ_s by

$$\hat{\lambda}_s = \lambda_s \varepsilon_s^* = \lambda_g \prod_{j=1}^{K} A(j,s(j)) \cdot \varepsilon_s^* \quad , \tag{4}$$

where ε_s^* is a random error term. Upon taking logarithms of Eq. (4), we obtain the model given by

$$\hat{\tau}_s = \tau_s + \varepsilon_s = \tau_g + \sum_{j=1}^{K} B(j,s(j)) + \varepsilon_s \quad , \tag{5}$$

where $\hat{\tau}_s = \log(\hat{\lambda}_s)$, $\tau_s = \log(\lambda_s)$, $\tau_g = \log(\lambda_g)$, $\varepsilon_s = \log(\varepsilon_s^*)$, and $B(j,i(j)) = \log[A(j,s(j))]$. Thus, Eq. (5) is a linear model in the logarithms of the factor effects (a so-called log-linear model), a fact that is important in obtaining the required failure rate estimates.

It is further assumed that $\hat{\tau}_s$ is an approximately normally distributed random variable. A small-scale Monte Carlo simulation was conducted to examine the appropriateness of this assumption. In the simulation, 1000 Poisson samples with mean $\theta = \lambda T$ were obtained. The empirical cumulative distribution function (cdf) of the 1000 samples was computed, as well as 95% confidence bounds on the empirical cdf. These results were graphically drawn on an overlay plot of the corresponding normal cdf computed from the first two sample moments of the 1000 samples. A typical plot is shown in Fig. 1 in which $\theta = 5$. It is observed that the assumption is not too bad when $\theta = 5$. The accuracy of the approximation increases (that is, $\hat{\tau}$ becomes more normally distributed) as θ increases and is excellent when θ is at least 10. When θ is as small as 1 or 2, the assumption is not unreasonable, although it is significantly worse than when $\theta = 5$. This is primarily due to the fact that when $\theta = 1$ or 2, $\hat{\tau}$ assumes a much smaller set of likely (probable) discrete values than when $\theta = 5$ and the use of a continuous normal distribution as an approximation is less satisfactory. In general, though, the assumption is reasonable and is likely to be compatible with the quality of the input data, as well as with the results.

The factors F_j may be either fixed quantities or random variables. Martz and Beckman[1] discuss the important differences in the interpretation of fixed and random factors, their impacts on the underlying statistical populations, and the interpretation of the statistical inferences about λ.

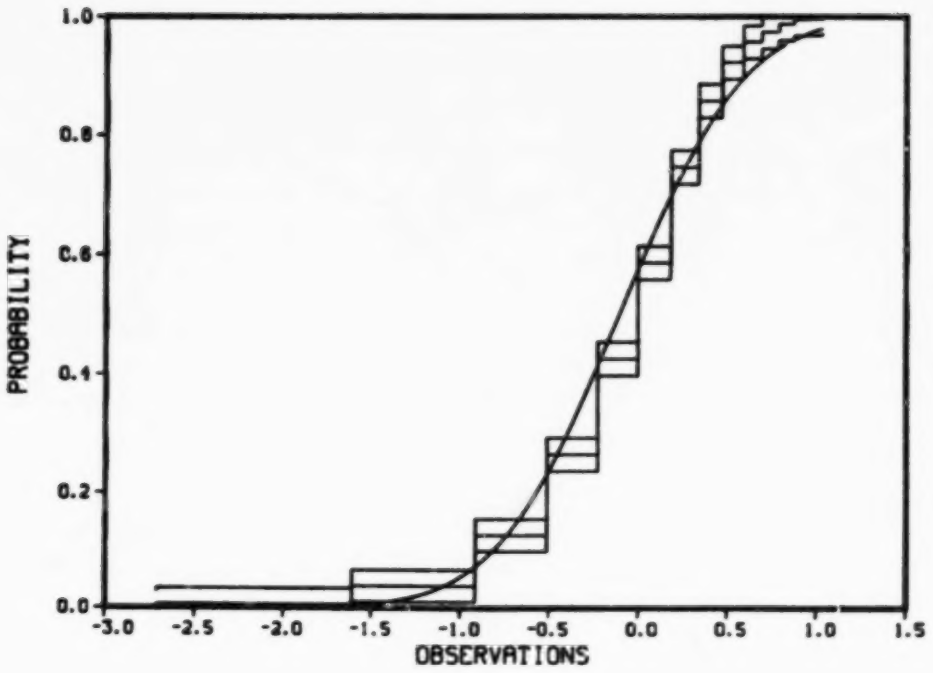


Fig. 1.
A comparison of the empirical distribution of $\bar{\tau}$ with a normal distribution when $\theta = 5$.

If F_j is a random factor, then the corresponding log effect $B(j,s(j))$ is assumed a priori to be the value of a normally distributed random variable, which implies that $A(j,s(j))$ has a log-normal distribution.

If we let I_1 denote the set of indices of the fixed effects in Eq. (5) and let I_2 denote the set of indices corresponding to the random effects, Eq. (5) may be written as

$$\bar{\tau}_s = \tau_g + \sum_{j\epsilon I_1} B(j,s(j)) + \sum_{j\epsilon I_2} B(j,s(j)) + \epsilon_s \quad . \tag{6}$$

Also,

$$\tau_s = \tau_g + \sum_{j\epsilon I_1} B(j,s(j)) + \sum_{j\epsilon I_2} B(j,s(j)) \quad . \tag{7}$$

By propagation of errors it is easily shown that $\mathrm{Var}(\bar{\tau}_s) \sim 1/E(\lambda_s)T_s$. Thus, it is assumed that ϵ_s is normally distributed with mean 0 and variance

$\sigma_\varepsilon^2/E(\lambda_s)T_s$. It follows that ε_s^* has a log-normal distribution with median 1, mean $\exp(\sigma_\varepsilon^2/2E(\lambda_s)T_s)$, and variance $\exp(2\sigma_\varepsilon^2/E(\lambda_s)T_s) - \exp(\sigma_\varepsilon^2/E(\lambda_s)T_s)$.

Since Var $(\hat{\tau}_s) \sim 1/E(\lambda_s)T_s$, we further assume that each random effect B(j,s(j)) has a normal distribution with mean zero and variance $\sigma_j^2/E(\lambda_s)T_s$ and, further, that all random effects, as well as ε_s, are independently distributed. Thus, $\hat{\tau}_s$ is normally distributed with mean

$$\tau_g + \sum_{j\in I_1} B(j,s(j))$$

and variance

$$\frac{\sum_{j\in I_2} \sigma_j^2 + \sigma_\varepsilon^2}{E(\lambda_s)T_s} \quad .$$

For example, in the previous pump example let us consider only two factors F_1 (plant ID) and F_2 (system ID). Further, suppose that only two plants are reported in the data base. Thus, it may be appropriate to consider the plant factor as a random effect from the total plant population. Further, suppose that all 45 levels of "system ID" are fixed effects in the sense that these levels represent all possible plant systems in which pumps could possibly be used. In this case, the model for the logarithm of the estimated pump failure rate would be

$$\hat{\tau}_s = \tau_g + B(1,s(1)) + B(2,s(2)) + \varepsilon_s \quad ,$$

where $B(1,s(1)) \sim N(0,\sigma_1^2/E(\lambda_s)T_s)$ and the values B(2,s(2)) are constants. If there is no plant-to-plant contribution to the estimated log-failure rate, then $\sigma_1^2 = 0$ and $B(1,s(1)) \equiv 0$. This corresponds to a multiplicative plant-to-plant effect of unity, that is, $A(1,s(1)) \equiv 1$.

In addition, the fixed factors are assumed to have corresponding effects that sum to 0 over all the levels of that factor. That is, if F_j is a fixed factor having m_j levels, then

$$\sum_{s(j)=1}^{m_j} B(j,s(j)) = 0 \quad .$$

This assumption is quite common in statistical analysis of variance where it is used to allow various design matrices to be of full rank. Thus, in the above example, the system ID factor is assumed to have corresponding effects such that

$$\sum_{s(2)=1}^{45} B(2,s(2)) = 0 \quad .$$

If there is <u>no</u> system effect for any of the 45 system applications, then $B(2,s(2)) \equiv 0$, $s(2) = 1, 2, \ldots, 45$ and the assumption is trivially true.

From the above assumptions, it follows that

$$\tau_s \sim N\left[\tau_g + \sum_{j\epsilon I_1} B(j,s(j)), \sum_{j\epsilon I_2} \sigma_j^2/E(\lambda_s)T_s\right] \qquad (8)$$

in those cases in which there are one or more random factors present in the analysis (that is, $I_2 \neq \phi$). Since τ_s is a random variable, it follows that the failure rate λ_s is likewise a random variable having a log-normal (prior) distribution. Both point and confidence interval estimates may be found for the mean of the distribution in Eq. (8).

It also follows that

$$\hat{\tau}_s \sim N\left[\tau_g + \sum_{j\epsilon I_1} B(j,s(j)), \frac{\sum_{j\epsilon I_2} \sigma_j^2 + \sigma_\epsilon^2}{E(\lambda_s)T_s}\right] ,$$

where $\hat{\tau}_s = \log(r_s/T_s)$.

In order to stabilize the variance in this equation, consider the transformation of the $\hat{\tau}_s$ values defined by

$$\hat{\eta}_s = \sqrt{E(\lambda_s)T_s}\ \hat{\tau}_s \ .$$

Thus,

$$\hat{\eta}_s \sim N\left\{\sqrt{E(\lambda_s)T_s}\left[\tau_g + \sum_{j\in I_1} B(j,s(j))\right] \ , \ \sum_{j\in I_2} \sigma_j^2 + \sigma_e^2\right\} \ . \qquad (9)$$

In practice in computing $\hat{\eta}_s$, r_s, the observed number of failures in cell s, is used as an estimate of the unknown value $E(\lambda_s)T_s$.

Tolerance interval estimates can be computed for both the population of τ_s values (when there is at least one random factor in the analysis) and the population of $\hat{\tau}_s$ values for any given time interval. From these intervals, corresponding tolerance intervals on either λ_s or $\hat{\lambda}_s$ can likewise be computed. The mathematical details for constructing such estimates will be considered in Sec. III.B.

Based on the above normality assumptions and log-linear model, FRAC computes estimates for the unknown mean and variance in Eq. (9) for any specified cell s. The method of weighted maximum likelihood (ML) or weighted restricted maximum likelihood (REML) is used to estimate the unknown parameters for the case in which there is at least one random factor present. In the case in which there are no random factors present, ordinary weighted least squares method is used to obtain the required estimates of the fixed effects. A complete description of the ML and REML methods and algorithms used to obtain the estimates for the unbalanced mixed effects model (the model and situation specialized here) is given by Hartley and Rao,[6] Jennrich and Sampson,[7] and Harville.[8] In addition, FRAC computes the variances and covariances of the estimates that are used in constructing the confidence and tolerance interval estimates.

III. FRAC APPLICATION USER'S GUIDE

The purpose of this section is to provide the reliability analyst with information on how to use the FRAC computer code for conducting failure rate analysis of variance. Primary considerations include a description of the input requirements, format, and procedures, and a discussion of the output,

which includes the corresponding statistical formulas, as well as their use and interpretation.

A. FRAC Input Description

The input used to set up FRAC consists of the following information:

- TITLE identification (optional)
- the MODEL specified for use in the analysis
- the ANAlysis upon which the estimation procedure is based (optional)
- the DATA upon which the analysis is based
- TOLerance interval coverages and confidences desired
  - a set of user-specified cells for which failure rate estimates are desired (optional)
- a HYPothesis to be tested regarding which factors significantly affect the failure rate (optional)
- the END of input data.

The above information is entered into FRAC in a series of so-called "paragraphs" known as a title paragraph, model paragraph, analysis paragraph, data paragraph, tolerance interval paragraph, hypothesis paragraph, and end paragraph. A slash (/) encountered as the first entry in a line denotes a new paragraph. Within a paragraph the required input information is structured as described for each paragraph below. Also, paragraphs should be input in the order listed below. The word "PARAGRAPH" may either be typed or not typed as part of the identification on line 1 of each paragraph described below.

1. /TITLE Paragraph (Optional).

(a) /TITLE on line 1 (no restrictions on spacing)

(b) Up to 20 lines of title information (anything the user desires to use as header ID information on the output), beginning on line 2 of the TITLE paragraph.

2. MODEL Paragraph.

(a) /MODEL on line 1

(b) Up to 20 lines of free-format model specifications. Each line is free-format with information entered in the following order.

(1) Factor name (up to 8 letters and/or characters).

(2) Factor type. Type in F or FIXED for fixed factors. Type in R or RANDOM for random factors.

NOTE 1: If any main or nested factor in an interaction is a random factor, then the interaction factor itself is a random factor.

(3) Maximum number of levels if the factor is either a fixed main or nested factor (if the factor is a random factor, there is no need to enter the maximum number of levels; if the factor is an interaction factor, there is also no need to enter the maximum number of levels).

NOTE 2: The maximum number of levels of a factor will usually simply be the number of levels of that factor. The exception occurs when the factor is nested and the number of levels depends upon the particular level of the factor in which it is nested. For example, in the hierarchical pump design breakdown given in Sec. I, the primary factor has 2 levels while the secondary factor (which is nested with the primary factor) has 3 levels for a positive displacement pump design and only 2 levels for a centrifugal pump design. Thus, the maximum number of levels for the secondary design breakdown factor is 3.

(4) Integer factor subscript(s) that identify the factor as a non-nested main factor (single subscript required), a nested main factor (multiple subscripts required), or an interaction (multiple subscripts required). Delimit multiple subscripts by either blanks or commas.

The rules for assigning subscripts follow the general conventions used in statistical analysis of variance and are summarized as follows:

- Main factors are specified first followed by any interaction factors that are included in the model.
- Main non-nested factors each have a single subscript that reflects the order in which the factor appears in the input data cell (see DATA paragraph below). The single subscript is also the order in which the factor appears in the MODEL paragraph.
- Main nested factors have multiple subscripts that include all the subscripts of the factor in which it is nested followed by a unique integer value that reflects the order in which the nested factor appears in the input data cell (see DATA paragraph below). The unique integer value is also the order of the nested factor in the MODEL paragraph.

- Interaction factors have multiple subscripts (which are subsequently converted to a single subscript by FRAC) which must include the subscripts of all the factors contained in the interaction. In the case of an interaction, there is no unique integer subscript value associated with the interaction, as in the case of main nested factors. Such a unique identifier is automatically assigned by FRAC.

- Multiple subscripts must appear in increasing order from left to right.

Two examples on the use of the MODEL paragraph will now be given as an aid in understanding the use of these rules.

Example 1: Pump Failure Rate Analysis 1

```
/MODEL
 PLANT          RANDOM                 1
 SYSW/PLT       RANDOM                 1,2
 TYPE           FIXED          2       3
 HEAD           FIXED          4       4
 HSPWR          FIXED          4       5
 TYP*HEAD       FIXED                  3,4
 TYP*SYS        RANDOM                 1,2,3
```

In this example, PLANT is a random factor having subscript 1; thus, F_1 is a random plant factor in the analysis. SYSW/PLT is a random system-use factor that is nested within the PLANT factor in this analysis. That is, the systems identified in which pumps are used differ for each of the plants considered in the analysis due to different plant designs, etc. Actually, five systems were randomly selected from each of two plants used in the study in order to reduce the computational burden of analyzing all 45 systems. Thus, F_2 is the system factor. Since this is a nested factor, we have the subscripts 1,2 associated with it. F_3 is the fixed primary pump TYPE factor having two levels, positive displacement [s(3) = 1], and centrifugal [s(3) = 2]. The head factor, HEAD, F_4 is fixed at 4 levels; namely, 1-10, $10\text{-}10^2$, $10^2\text{-}10^3$, and $> 10^3$ (psig), denoted by s(4) = 1,2,3,4, respectively. The pump horsepower factor, HSPWR, F_5 is also fixed at 4 levels; namely, < 1, 1-10, $10\text{-}10^2$, and $> 10^2$ (hp). These are likewise denoted by s(5) = 1,2,3,4, respectively. The interaction factor, TYP*HEAD, represents a possible interaction between pump type and head size in the sense that the effect of head size on pump failure rate may depend on the type of pump, that is, whether the pump is a positive displacement or a centrifugal design.

Since this factor is an interaction between main factors F_3 and F_4, the subscripts on this interaction (which is interpreted automatically as factor F_6 by FRAC) are 3,4. The TYP*SYS factor represents a possible interaction between the fixed pump type factor F_3 and the random nested system factor F_2. Since F_2 is a random factor, this interaction (automatically denoted as factor F_7 by FRAC) is also a random factor. The corresponding subscripts are 1,2,3 since F_2 has subscripts 1,2 associated with it and F_3 has subscript 3 associated with it. Note that all multiple subscripts are given in increasing order; however, they do not have to be consecutive integers as they are in this example. For example, if an interaction between pump type (F_3) and horsepower (F_5) were included in the model, then this interaction factor would have subscripts 3,5.

Example 2: Pump Failure Rate Analysis 2

```
/MODEL
 PLANT        RANDOM                 1
 SYSTEM       FIXED          10      2
 PTYPE        FIXED          2       3
 STYPE        FIXED          3       3,4
 TTYPE        FIXED          4       3,4,5
 HEAD         FIXED          4       6
 PT*HD        FIXED                  3,6
 ST*HD        FIXED                  3,4,6
 PL*PT*HD     RANDOM                 1,3,6
```

In this example, PLANT (factor F_1) is again a random factor since only two plants out of the population of all existing commercial US nuclear power plants were included in the data base. Here 10 specific systems are being analyzed for their effect on pump failure rates; thus, F_2 is a fixed factor having 10 levels. According to the hierarchical pump design classification given in Sec. I, PTYPE (factor F_3) is the primary classification factor fixed at 2 levels; STYPE (factor F_4) is the nested secondary classification factor fixed at a maximum of 3 levels; TTYPE (factor F_5) is a nested tertiary factor fixed at a maximum of 4 levels. HEAD (factor F_6) is fixed at 4 levels as in Example 1. PT*HD (factor F_7) represents a possible interaction between factors F_3 and F_6; ST*HD (factor F_8) represents a possible interaction between factors F_4 and F_6 (since nested factor F_4 has subscripts 3,4 this interaction must likewise have subscripts 3,4,as well as 6). Finally, PL*PT*HD (factor F_9) represents a possible three-factor

interaction between PLANT (factor F_1), PTYPE (factor F_3), and HEAD (factor F_6). Thus, the subscripts on this interaction are 1,3,6.

3. ANAlysis Paragraph (Optional)

(a) /ANAlysis on line 1, where the characters ANA must be typed and the remainder of the word, "lysis," may or may not be typed, as desired.

(b) Values of IREML and IVARCOV in free-format on line 2. The value IREML = 0 indicates that ordinary (unrestricted) maximum likelihood (ML) estimates are to be used in the analysis while IREML = 1 indicates that restricted maximum likelihood (REML) estimates are to be used.[7,8] Also, IVARCOV = 1 indicates that a Fisher scoring algorithm is used in the analysis while IVARCOV = 2 indicates the choice of a Newton-Raphson algorithm.[7]

NOTE 3: If the ANAlysis paragraph is missing, the default values used in the analysis are IREML = 0 and IVARCOV = 1.

4. DATA Paragraph

(a) /DATA on line 1

(b) Location/format specification on line 2. If input failure data are to be included in the data paragraph itself, give the format specification in single quotes; for example, '(3I4,F12.6)', '(5I2,E12.4)', 'FREE', etc. If the failure data to be analyzed are located in an external file, give the file name (in single quotes) first followed by the format specification for reading the file; for example, 'PUMP2.DAT','(6I4,F12.3)'.

FRAC assumes that the input reliability data are ordered as follows:

(1) Row vector of (integer) levels of each of the main factors in the same order as specified in the MODEL paragraph.

NOTE 4: The levels of the interaction factors are not included because these are automatically computed by FRAC.

NOTE 5: If the levels are read by means of a 'FREE' format specification, then delimit the levels of each factor by either blanks or commas.

NOTE 6: The level corresponding to a fixed factor F_j is a value of the integer s(j) discussed in Sec. II and must be an integer from the set $\{1, 2, \ldots, m_j\}$, where m_j is the maximum number of levels specified in the MODEL paragraph.

NOTE 7: The levels corresponding to a random factor F_j do not have to be integers from the set $\{1, 2, \ldots, m_j\}$. For example, the two levels of the random PLANT factor F_1 in Example 1 could be coded in the first column of the input failure data as 5 and 10 corresponding to, say, plants having IDs equal to 5 and 10, respectively. The same analysis results will be obtained if any two integers are used to distinguish the plant to which the pump failure data pertains since PLANT (F_1) is a random factor.

(2) Number of failures (an integer variable) (r in Sec. I).

(3) Total observed operational/test time (a real variable) (T in Sec. I).

(c) If input reliability data are included in the DATA paragraph, insert the data here.

5. TOLerance Interval Paragraph

(a) /TOL on line 1

(b) Values of IND, ALPHA, BETA, and TO in free-format on line 2.

NOTE 8: IND = 1 if tolerance interval estimates are desired on the failure rate population given in Eq. (8); IND = 0 if tolerance interval estimates are desired on the population of failure rate estimates given by Eq. (9).

NOTE 9: 100(1 - ALPHA)% - 100(1 - BETA)% tolerance interval estimates are computed, where 100(1 - ALPHA)% is the "confidence" that the interval contains at least 100(1 - BETA)% (the "coverage") of the underlying population (see Sec. III.B).

NOTE 10: TO is the specified operating time interval for which tolerance interval estimates are to be computed. If TO is set to 0, the default value of TO = 8760 is used in the analysis. TO (a real variable) must be specified in hours.

(c) On lines 3, 4, ... , type the factor levels of all main factors corresponding to all cells (one per line) specified by the user for which failure rate estimates are desired. If none are desired, omit (c). Type a zero for the factor level corresponding to all random main factors. If the user wishes to specify weights (which must sum to 1) for each cell specified, they must be typed after the vector of factor levels.

NOTE 11: The input data in (c) are useful for estimating failure rates for any desired combination of factor levels, even for those for which no reliability data is available.

Example 3

In Example 1, suppose that the analyst desires a tolerance interval estimate on the population of pump failure rates for a 100-hp centrifugal pump having a head of 50 psi, even though no pumps having this description are reported in the reliability data used in the analysis. In this case, the vector of levels (the cell) denoted by 0,0,2,2,3 would be input in (c) above, and the output would be the desired interval estimate for the stated combination of factor levels.

NOTE 12: The input data in (c) are also useful for estimating component failure rates in which not all of the levels of the factors are known (see Sec. I). In this case, the user specifies in (c) all of the likely (unknown) cells and a possible weight (the likelihood) of each cell. If all specified (unknown) cells are thought to be equally likely, then weights do not need to be given for the cells in (c). FRAC produces tolerance interval estimates for each (unknown) cell, as well as various kinds of averages over all the cells specified in (c). These averages are then used as the desired estimates for the unknown use conditions (a mathematical description of these averages is given in Sec. III.B.18).

Example 4

Again let us consider Example 1. Suppose that a plant is to be built in which an estimate is required of the failure rate of a 100-hp pump having a 50-psi head. Thus, it is appropriate to "average" over both types of pumps (factor F_3), since the type of pump has not been identified. This is accomplished by specifying the two (unknown) cells given by 0,0,1,2,3 and 0,0,2,2,3 in (c). If both pump types are believed to be equally likely, no weights would be used. However, if a centrifugal pump is thought to be roughly twice as likely as a positive displacement design, then the information in (c) above would be typed in on two lines as follows:

```
0,0,1,2,3      .33
0,0,2,2,3      .67
```

In this case, weighted averages over both of these cells would be taken using the indicated weights.

NOTE 13: Generally speaking, the fewer factor levels that are known, the wider will be the average tolerance interval failure rate estimates. That is, as the number of (unknown) cells in (c) increases, the wider will be the corresponding average tolerance interval estimates obtained by averaging over the (unknown) cells. In other words, the more that is known about the operating/design/environment/maintenance conditions associated with a component such as a pump, the narrower the failure rate uncertainty bounds (the width of the tolerance interval estimate) associated with the component.

6. HYPothesis Paragraph.

The HYPothesis paragraph is unavailable in the current version of FRAC. It will be incorporated in a future version of FRAC as follows:

(a) /HYP on line 1

(b) Suppose we are interested in testing k hypotheses, where $k \geq 1$. Beginning on line 2, type the number of the hypothesis and the name (or names) of the factor (or factors) corresponding to the hypothesis to be tested. That is, specify the name(s) of the factor(s) to be excluded from the model in the hypothesis test. The corresponding difference between the value of -2 log (maximum likelihood) for this reduced model and that of the full (unreduced) model is a likelihood ratio test statistic of the factors that have been set to zero in the model (that is, the stated hypothesis).

(c) Repeat (b) for each of the k hypotheses.

Example 5

In Example 1, suppose that the following HYPothesis paragraph is specified.

```
/HYP
1PLANT
2TYP*HEAD
3HEAD TYPE
```

In hypothesis number 1, the single main factor PLANT is set to zero and is used to examine the statistical significance of this factor in the model. In the second hypothesis, the single interaction factor TYP*HEAD is to be tested for statistical significance. The third hypothesis gives a statistical test of the joint significance of both the main factors HEAD and TYPE in the model.

NOTE 14: A user can separately test the statistical significance of each and every factor in the model by listing the name of each and every factor in the model (see MODEL paragraph) on a separate line in the HYPothesis paragraph.

7. END Paragraph

(a) /END on line 1 (signifies the end of all input data)

8. Examples. The following examples illustrate the structure of the various input paragraphs described above.

Example 6: Failure Data Included in DATA Paragraph

```
/TITLE  PARAGRAPH
    TEST ROUTINE WITH SIMULATED DATA NOVEMBER 1981
/MODEL PARAGRAPH
PLANT    RANDOM          1
SYSW/PLT    RANDOM       1,2
TYPE  FIXED  2           3
```

```
HEAD  FIXED  4                4
HSPWR     FIXED  4              5
TYP*HEAD FIXED                 3,4
TYPE*SYS RANDOM                1,2,3
/ANALYSIS PARAGRAPH
          0          1
/DATA PARAGRAPH
DATA IS INCLUDED IN INPUT FILE, FORMAT IS 'FREE'
1,1,1,3,4  37  .14E+06
1,1,1,4,3  3  .175E+05
1,1,1,4,4  8  .175E+05
1,1,2,2,2  2   .175E+05
1,1,2,3,3  5   .175E+05
1,1,2,3,4 31   .105E+05
1,1,2,4,4 11   .350E+05
1,2,1,2,2  3   .175E+05
1,2,1,2,3  28  .701E+05
1,2,2,1,1  0   .175E+05
1,2,2,1,2  3   .175E+05
1,2,2,1,3  29   .526E+05
1,2,2,2,2  27   .701E+05
1,2,2,3,4  16   .175E+05
2,1,1,3,2   10   .175E+05
2,1,1,3,3  25  .175E+05
2,1,2,1,2   109   .526E+05
2,1,2,2,2   43    .175E+05
2,1,2,2,3   1058  .298E+06
2,1,2,2,4   198   .350E+05
2,2,1,3,2   64,   .701E+05
2,2,1,3,3  42  .350E+05
2,2,1,4,2   39   .35E+05
2,2,2,4,2   36   .35E+05
/TOLERANCE PARAGRAPH
         1   5.0000000000000000E-02   0.1000000000000000   0
0,0,1,3,3  .2
0,0,2,1,4  .4
0,0,1,2,4  .3
0,0,2,1,3  .1
/END
```

Example 7: Failure Data in External File

```
/TITLE PARAGRAPH
    SAMPLE RUN TO TEST ROUTINE WITH SIMULATED DATA
/MODEL PARAGRAPH
PLANT  RANDOM           1
SYSW/PLT  RANDOM        1,2
TYPE  FIXED    2        3
HEAD  FIXED    4        4
HSPWR  FIXED    4       5
TYP*HEAD  FIXED         3,4
TYP*SYS  RANDOM         1,2,3
/ANALYSIS PARAGRAPH
         0       1
```

```
/DATA PARAGRAPH
DATA IS ON FILE 'PUMP3.DAT' USING FORMAT '(6I4,F12.4)'
/TOLERANCE PARAGRAPH
0        .05       .10       0
0,0,2,1,1           .4
0,0,2,3,1           .5
0,0,2,4,1           .1
/END
```

The external file named 'PUMP3.DAT' contains the following failure data.

```
1    1    1    3    4    37   .14E+06
1    1    1    4    3    3    .175E+05
1    1    1    4    4    8    .175E+05
1    1    2    2    2    2    .175E+05
1    1    2    3    3    5    .175E+05
1    1    2    3    4    31   .105E+06
1    1    2    4    4    11   .350E+05
1    2    1    2    2    3    .175E+05
1    2    1    2    3    28   .701E+05
1    2    2    1    1    0    .175E+05
1    2    2    1    2    3    .175E+05
1    2    2    1    3    29   .526E+05
1    2    2    2    2    27   .701E+05
1    2    2    3    4    16   .175E+05
2    1    1    3    2    10   .175E+05
2    1    1    3    3    25   .175E+05
2    1    2    1    2    109  .526E+05
2    1    2    2    2    43   .175E+05
2    1    2    2    3    1058 .298E+06
2    1    2    2    4    198  .350E+05
2    2    1    3    2    64   .701E+05
2    2    1    3    3    42   .350E+05
2    2    1    4    2    39   .350E+05
2    2    2    4    2    36   .350E+05
```

B. Description and Discussion of FRAC Output

A sample of the output of FRAC follows this subsection. The entire listing is presented, and each portion of the output is numbered from 1 to 21. For each of these portions of the output there is a corresponding subsection numbered 1 to 21, which describes the output.

The sample output is based on a set of simulated pump failure data given in Example 5. Recall that each pump failure can be assigned to a cell. The cells comprise random factors due to the plants and the systems within plants in which the pumps were used. In addition, each cell is characterized by three fixed factors: the type of pump, the head size, and the horsepower. Two additional factors were assumed, a fixed effect type by head size interaction and a random type by system interaction.

1. Input Data. The input data are listed.

2. LN Failure Rate Analysis Model. Each of the main factors and interactions is listed along with its factor index, type, and subscripts (see Sec. III.A.2). In addition, the number of levels for each factor is listed.

In the example, there are five main factors and two interactive factors. The number of levels of the fixed interactive factors is determined by the product of the levels of the corresponding main factors.

In addition, the total number of cells read is also computed and printed.

3. Failure Rate Information by Cells. For each cell for which there is data, the cell index (labeled "i" in Sec. II), along with the cell factor levels, is listed. The number of failures R and the time on test T are also given. The failure rate is R/T, and LN(F.R.) stands for the natural logarithm of the failure rate, log(R/T). Using the results of Welker and Lipow,[9] R is assumed to be 1/3 when there are no failures. The weight is the inverse of the variance of log(R/T), which is derived by propagation of error as

$$\mathrm{Var}[\log(R/T)] \sim \left[\frac{\partial \log(R/T)}{\partial R}\right]^2 \mathrm{Var}(R) \quad .$$

Since R is assumed to be a Poisson variable, $\mathrm{Var}(R) = E(\lambda_s)T_s \sim R$, and

$$\left[\frac{\partial \log(R/T)}{\partial R}\right]^2 = \left(\frac{1}{R^2}\right) \quad .$$

Therefore, $\mathrm{Var}[\log(R/T)] \sim 1/R$ and the weight assigned to each cell is R.

In the example given here, the cell index gives the levels for the PLANT, SYSTEM within plant, TYPE, HEAD, HORSEPOWER, the TYPE by HEAD interaction, and the TYPE by SYSTEM interaction. Therefore, for data points with index number 1, the cell is the one at which the level of the PLANT is 1, SYSTEM within plant is 1, the TYPE is level 1, the HEAD is level 3, the HORSEPOWER is level 4, the level of the TYPE by HEAD interaction is 3, and that of the TYPE by SYSTEM interaction is 1. The levels of the interactions are derived from the levels of the main factors such that every combination of the main factors in the interaction (for example, the 8 combinations of TYPE and HEAD levels) is assigned a unique level, as discussed in Sec. III.A.2.

4. Estimates of Fixed Effects. The levels of the fixed effects among the $B(j,s(j)) = \log(A(j,s(j)))$ are estimated by the method of maximum likelihood or restricted maximum likelihood, as stated in Sec. II. Since, by assumptions of Sec. II, the sum of all of the levels of a fixed factor is zero, estimates are given for only the logarithms of the first $(m_j - 1)$ levels of the fixed effect. The last level of the logarithm of any fixed effect may be estimated by

$$-\sum_{s(j)=1}^{m_j - 1} \hat{B}(j,s(j)) \quad ,$$

where $\hat{B}(j,s(j))$ is the estimated effect of the logarithm of the j^{th} factor at the $s(j)^{th}$ level. In the example given here, the main effect TYPE is at two levels. Therefore, the second level of TYPE is estimated as $\hat{B}(3,2) = -\hat{B}(3,1) = 0.2463$, which is the log effect due to a centrifugal-type pump design.

The standard deviation of the estimated logarithms of the fixed effects is given by maximum likelihood or restricted maximum likelihood. Estimates of zero for these standard deviations indicate a possible over-parameterization of the model. This means that not enough cells were represented by data. In these cases, the estimated effects given by FRAC may not be the same as those given by BMDP3V.[10] However, estimates of all estimable functions (linear combinations of the estimates of the fixed effects that correspond to linear combinations of the design matrix of Sec. III.B.9) will be identical to BMDP3V.

The standard deviation of the logarithm of the remaining (unstated) level of a fixed effect may be obtained by taking the square root of the quantity

$$\mathrm{Var}\left[-\sum_{s(j)=1}^{m_j-1} \hat{B}(j,s(j))\right] = \sum_{s(j)=1}^{m_j-1} \mathrm{Var}(\hat{B}(j,s(j)) + \sum_{s(j)\neq s(k)}^{m_j-1} \mathrm{Cov}(\hat{B}(j,s(j)),\hat{B}(j,s(k))).$$

The values of the variances and covariances are discussed in Sec. III.B.7. In the problem shown here, the standard deviation of the estimate of the logarithm of the second level of the factor TYPE is given by 0.04874.

The estimates of the levels of the fixed effects $\hat{A}(j,s(j))$ are taken as $\hat{A}(j,s(j)) = \exp(\hat{B}(j,s(j)))$. The standard deviations of $\hat{A}(j,s(j))$ are found by propagation of error to be the square root of the quantity

$$\mathrm{Var}(\hat{A}(j,s(j))) \sim \exp(2\hat{B}(j,s(j))) \cdot \mathrm{Var}(\hat{B}(j,s(j))) \quad .$$

In the example, the estimated effect for the first level of the factor TYPE (that is, a positive displacement pump) is given by $\hat{A}(2,1) = \exp(\hat{B}(2,1)) = \exp(-0.2463) = 0.7817$. The standard deviation of this estimate is approximately given by $\exp(-0.2463) \cdot (0.04874) = 0.03810$.

5. Random Effects Variance Estimates. As indicated in Sec. II, the logarithm of each of the random effects is assumed to have mean zero. In addition, the logarithm of each of these effects is also assumed to have a normal distribution with an unknown variance $\sigma_j^2/E(\lambda_s)T_s$. The parameters, σ_j^2, are estimated by the methods of maximum likelihood or restricted maximum likelihood. For purposes here, we denote the estimate of σ_j^2 of the j^{th} random factor as $\hat{\sigma}_j^2$. In the example given here, the estimate of σ_j^2 for the logarithm of the plant effect is 0.9762.

The standard deviations of these estimates are given in the column marked STD DEV. The estimate divided by its standard deviations is given in the column EST/(STD DEV). In the example shown here, the standard deviation of the estimated σ_j^2 for the effect plant is 0.9873, while the ratio of the estimate to its standard deviation is 0.9888.

6. -2 log(Maximum Likelihood). This quantity is the likelihood function evaluated at the maximum likelihood estimates of the parameters. The hypothesis of no effect of a factor may be tested by taking the difference between this statistic when the model without the tested factor is fitted to the data and the statistic obtained by fitting the full model. If the factor tested has no effect, this difference should have an approximate chi-square (χ^2) distribution with d degrees of freedom. For fixed factors, d equals the number of levels fitted, which is the number of levels minus one for the main effects, while d = 1 for random factors.

In the example given here, the random effect plant could be tested for significance by fitting the data to the model without the plant effect and taking the difference between the two values of -2 log(maximum likelihood). In this case, if there is no plant effect, the resulting value should have an

approximate χ^2 distribution with one degree of freedom. If the difference exceeds, say, the 95th percentile of the χ^2 distribution, then the PLANT effect is significant at the 5% level of significance and thus belongs in the model. To test for significance of the fixed effect HEAD, the resulting χ^2 distribution would have three degrees of freedom.

7. Variance-Covariance Matrix for Fixed Effects. The symmetric matrix of numbers given here represents a listing of the estimated variances and covariances of the estimated logarithms of the levels of the fixed effects. The values given on the diagonal are the variances corresponding to each of the levels of the variables, while those off the diagonal are the covariances. For example, the variance of the first level of the factor TYPE is given by 0.002376, while the covariance between this variable and the second level of the factor HEAD is given by 0.00145. These estimates are asymptotic estimates, which means that they are the estimates of the variances and covariances of the estimated logarithms of the levels based on a hypothetical infinite sample of data. These variances and covariances are obtained by the inversion of the estimated information matrix, which is described in the references on maximum likelihood given in Sec. II.

In the algorithm used to compute these estimated variances, the estimates are required to be non-negative. Estimated variances equal to zero indicate a possible over-parameterization of the problem in the sense that there are insufficient cells and cell combinations to be able to estimate all the required terms specified in the model.

It can be seen that the problem given here is over-parameterized; the variance of the estimated constant is estimated to be zero.

8. Variance-Covariance Matrix for Random Effects. As with the fixed effects, these quantities are the estimated asymptotic variances and covariances of the estimated parameters σ_j^2 of each of the random factors. Note that these estimated parameters of the random factors are independent of the estimated levels of the fixed factors, hence, the covariance between these two estimators is zero.

These estimated variances and covariances are also obtained as the inverse of the estimated information matrix. The values of the estimated variances are constrained to be non-negative, and zero estimates for the variances may indicate an over-parameterization of the problem.

In the problem here, the TYPE*SYS interaction factor has an estimated variance of zero. This indicates that too few data cells were available to estimate all of the unknown parameters.

9. Design Matrix (X). For each fixed factor in the analysis, dummy variables are generated that take on the values -1, 0, 1. These variables indicate the presence (1) or absence (0) of the level of the factor. Since levels of the fixed factors are assumed to sum to zero, only the number of levels minus one are represented for each main fixed factor. The last level of the fixed main factor is represented by all of the other levels with value -1. For interactive or nested factors, the values of the dummy variables are given by the multiplication of the appropriate dummy variables for the corresponding main effects. The first dummy variable represents the generic mean, and the remaining columns follow the same order as they appear in the model paragraph.

In this example, the factors occur in the following order: the constant or generic mean, TYPE at one level, the three levels of HEAD, the three levels of HSPWR, and the three levels of the TYP*HEAD interaction. Therefore, the cell with index equal to 1 comprises the generic failure rate effect, the effect due to the first level of TYPE, the third level of HEAD, and the fourth level of HSPWR since the horsepower factors are all at -1. The remaining three columns for the TYP*HEAD interaction are determined by the product of the coefficient value for TYPE and the three values for HEAD. In this case, the three interaction terms are 0, 0, 1.

Rows of the design matrix are used to estimate the mean value for any cell. This is accomplished by summing the products of the elements of the row of the design matrix and the corresponding estimates of the fixed effects. For example, using the first row of the design matrix, the estimated mean value of the log failure rate of the cell with index 1 is given by

$$\begin{aligned}(1)(-8.040) + (1)(-0.2463) + 0(0.0159) + 0(0.4452) + 1(-0.2186) \\ + (-1)(-2.515) + (-1)(0.3718) + (-1)(0.8336) + (0)(-0.3504) \\ + (0)(0.02895) + (1)(0.01806) = -7.1772 \ .\end{aligned}$$

10. LN(FAILURE RATE) Residual Analysis. As in Sec. II, let I_1 be the set of indices of the fixed factors and I_2 be the set of indices of random factors. Let the levels of the k factors associated with the s^{th} cell be

denoted by the set of pairs of indices $S = \{(1,s(1)),(2,s(2)), \ldots ,(k,s(k))\}$, where, for example, (2,s(2)) means that the second factor is at the s(2) level. Then from Eq. (8) in Sec. II, the logarithm of the s^{th} failure rate, τ_s, is assumed to have a normal distribution with mean value

$$\mu_s = \tau_g + \sum_{j \epsilon I_1} B(j,s(j))$$

and variance given by

$$\sigma_s^2 = \sum_{j \epsilon I_2} Var(B(j,s(j))) = \sum_{j \epsilon I_2} \sigma_j^2 \ .$$

Note that the mean value of the logarithm of the s^{th} failure rate is not a function of the random factors, and further, that the variance of this quantity is not a function of the fixed factors.

The PREDICTED LN FAILURE RATE given in the output is the estimated mean value for each OBSERVED LN FAILURE RATE. This estimated mean is given by

$$\hat{\mu}_s = \hat{\tau}_g + \sum_{j \epsilon I_1} \hat{B}(j,s(j)) \ ,$$

where $\hat{\tau}_g$ is the estimated constant in the model and $\hat{B}(j,s(j))$ is the estimate of the effect of the $s(j)^{th}$ level of the j^{th} fixed factor. Recall that if s(j) represents the last level of the j^{th} factor, then $\hat{B}(j,s(j))$ is estimated as the negative sum of the remaining factors.

The STD DEV PREDICTED is the square root of the estimated variance of $\hat{\mu}_s$. This variance is given by

$$Var(\hat{\mu}_s) = Var(\hat{\tau}_g) + \sum_{j \epsilon I_1} Var(\hat{B}(j,s(j)) + 2 \sum_{j \epsilon I_1} Cov(\hat{\tau}_g, \hat{B}(j,s(j)))$$

$$+ \sum_{j \epsilon I_1} \sum_{k \epsilon I_1} Cov(\hat{B}(j,s(j)), \hat{B}(k,s(k))) \ , \ j \neq k$$

where estimates of these variances and covariances are given in the variance-covariance matrix of the fixed effects (see Sec. III.B.7).

The 95% CONFIDENCE INTERVALS are approximate confidence intervals on the mean value μ_s of the logarithm of the failure rate. They are computed as

$$\hat{\mu}_s \pm 1.96 \sqrt{Var(\hat{\mu}_s)} \quad ,$$

where $\hat{\mu}_s$ and $Var(\hat{\mu}_s)$ are defined above. These intervals are interpreted to mean that approximately 95% of all such intervals will contain the true value of μ_s. Note that as the number of cells increases, the variability in the estimated value of μ_s decreases and the confidence intervals on μ_s become shorter. Therefore, these intervals should not be viewed as intervals containing a certain percentage of observed failure rates.

From Eq. (9) of Sec. II, the mean value of the logarithm of the observed failure rate is given by

$$\tau_s + \sum_{j \epsilon I_1} B(j,s(j)) \quad .$$

Therefore, the confidence interval just derived for the mean of the logarithm of the true failure rate is also a confidence interval on the mean of the logarithm of the observed failure rate.

The RESIDUAL (O-P) is the OBSERVED LN FAILURE RATE minus the PREDICTED LN FAILURE RATE. The STD DEV RESIDUAL is the square root of the variance of the RESIDUAL (O-P). This variance is estimated by adding the estimated variance of the observed logarithm of the failure rate to the estimated variance of the predicted logarithm of the failure rate. That is,

$$Var[\text{RESIDUAL (O-P)}] = Var(\hat{\mu}_s) + \Big(\sum_{j \epsilon I_2} \hat{\sigma}_j^2 \Big)/R_s \quad ,$$

where $Var(\hat{\mu}_s)$ and $\hat{\sigma}_s^2$ are defined above, and R_s is the observed number of failures in cell s. This estimator of the variance ignores the covariance between

the observed and the predicted logarithm of the failure rate and, as such, is only an approximation.

The STANDARDIZED RESIDUAL is the RESIDUAL (O-P) divided by the STD DEV RESIDUAL. This diagnostic quantity should give some idea of poorly fitted cells. If the absolute value of the standardized residual is larger than three or four, then the input data for this cell should be investigated.

11. Failure Rate Residual Analysis. The OBSERVED FAILURE RATE is the number of failures divided by the total time on test. The PREDICTED FAILURE RATE is the exponential of the PREDICTED LN FAILURE RATE of the previous section, and the limits on the 95% CONFIDENCE INTERVAL are the exponentials of the limits of the confidence intervals given in the previous section. The RESIDUAL (O-P) is the difference between the observed and predicted failure rates.

In the example, the predicted failure rate, 0.7639E-03, for the cell with index 1 is equal to exp(-7.177). The end points of the confidence interval are given by exp(-8.569) and exp(-5.785), respectively.

12. Tolerance Intervals on LN(FAILURE RATE) by Cells [Tolerance Intervals on LN(FAILURE RATE) Estimates by Cells]. In this section of the output, tolerance intervals are computed for each cell for which there are data. These intervals are on either the population of log failure rates, if there are random factors present, or the population of estimated log failure rates for a fixed time T_o.

A tolerance interval on the log failure rate or the logarithm of the estimated failure rate is an interval that includes at least a certain proportion, γ, of the statistical population of the log failure rate or the logarithm of the estimated failure rate at a specified confidence level $(1 - \alpha)$ for a time T_o. Such intervals are further defined and illustrated by Martz and Beckman.[10]

The first of these intervals is denoted by LN(FAILURE RATE) BY CELLS while the second is denoted by LN(FAILURE RATE) ESTIMATES BY CELLS. The possibility of the first type of tolerance interval arises from Eq. (8) where it is shown that the logarithm of the failure rates has a normal distribution with a variance that is the sum of all the variances of the random terms in the model scaled by $E(\lambda_s)T_o$. If the model has no random terms, then a tolerance interval cannot be computed, as there is no variability in the true underlying failure rate population.

The second type of tolerance interval, an interval on the population of failure rate estimates, may always be computed. From Eq. (9) of Sec. II, the logarithm of the observed failure rate has a normal distribution with a variance that is the sum of the variances of the random factors and the variance of the error term. The techniques for obtaining these two intervals are similar, and only the method for obtaining the intervals on the log failure rates will be given here.

To obtain tolerance intervals on the statistical population of the logarithm of the failure rate, the probability statement

$$\Pr\{\Pr\{L \leq \log \lambda \leq U\} \geq 1 - \beta\} \geq 1 - \alpha$$

must be solved for L and U, where γ and α are prespecified. Given values of L and U, we then state that we are at least $(1 - \alpha)\cdot 100\%$ confident that at least $(1 - \beta)\cdot 100\%$ of all of the values of $\log \lambda$ will be between L and U.

Approximate values of L and U are obtained in the following way. First, the $1 - \beta/2$ and $\beta/2$ percentiles of the distribution of the log failure rate over a time T_o are estimated. These estimates are given by

$$[\hat{\tau}_s + \sum_{j \in I_1} \hat{B}(j,s(j))] \pm Z_{1-\beta/2}[\sum_{j \in I_2} \hat{\sigma}_j^2/\hat{\lambda}_s T_o]^{1/2} \quad , \tag{10}$$

respectively, where

$$1 - \frac{\beta}{2} = \int_{-\infty}^{Z_{1-\beta/2}} \frac{1}{\sqrt{2\pi}} \exp(-t^2/2)dt \quad ,$$

and $\hat{\lambda}_s$ is estimated by

$$\hat{\lambda}_s = \exp\{\hat{\tau}_s + \sum_{j \in I_1} \hat{B}(j,s(j))\} \quad .$$

It follows that Eq. (10) is an estimator of the desired percentiles since, from Sec. II, log λ over time T_o was assumed to have a normal distribution with mean

$$\tau_s + \sum_{j \epsilon I_1} B(j,s(j))$$

and variance

$$\sum_{j \epsilon I_2} \sigma_j^2 / E(\lambda_s) T_o = \sum_{j \epsilon I_2} \mathrm{Var}(B(j,s(j))/E(\lambda_s) T_o \quad .$$

Letting A represent Eq. (10) with the positive square root, the value of U is estimated by

$$U = A + Z_{1-\alpha/2} [\mathrm{Var}(A)]^{1/2} \quad . \tag{11}$$

The quantity Var(A) is estimated by propagation of error to be

$$\begin{aligned} \mathrm{Var}(A) &= (A'(\hat{\tau}_s))^2 \mathrm{Var}(\hat{\tau}_s) \\ &+ \sum_{j \epsilon I_1} \sum_{k \epsilon I_1} A'(\hat{B}_j) A'(\hat{B}_k) \mathrm{Cov}(\hat{B}(j,s(j)) \, , \, \hat{B}(k,s(k))) \\ &+ \sum_{j \epsilon I_1} A'(\hat{\tau}_s) A'(\hat{B}_j) \mathrm{Cov}(\hat{\tau}_s, \hat{B}(j,s(j)) \\ &+ \sum_{j \epsilon I_2} \sum_{k \epsilon I_2} A'(\hat{\sigma}_j^2) \; A'(\hat{\sigma}_k^2) \mathrm{Cov}(\hat{\sigma}_j^2, \hat{\sigma}_k^2) \quad , \end{aligned} \tag{12}$$

where for any variable B, Cov(B,B) = Var(B),

$$2A'(\hat{\tau}_s) = 2 \frac{\partial A}{\partial \tau_s}$$

$$= 2 - Z_{1-\beta/2}\left[\sum_{j\in I_2} \frac{\hat{\sigma}_j^2}{T_0} \cdot \exp\{-\hat{\tau}_s - \sum_{j\in I_1} \hat{B}(j,s(j))\}\right]^{1/2} ,$$

$$2A'(\hat{B}_j) = 2\,\frac{\partial A}{\partial B(j,s(j))} = 2A'(\hat{\tau}_s) ,$$

and

$$2A'(\hat{\sigma}_j^2) = 2\,\frac{\partial A}{\partial \hat{\sigma}_j^2}$$

$$= Z_{1-\beta/2}\left[\exp\{-\hat{\tau}_s - \sum_{j\in I_1} \hat{B}(j,s(j))\}/(T_0 \cdot \sum_{j\in I_2} \hat{\sigma}_j^2)\right]^{1/2} .$$

All of the quantities in Eqs. (10)-(12) are estimated by using the estimates discussed in Sec. III.B.4-8.

The lower tolerance limit L is derived in a similar manner by letting C be Eq. (10) with the negative square root, and computing

$$L = C - Z_{1-\alpha/2}[\mathrm{Var}(C)]^{1/2} . \qquad (13)$$

In Eq. (13) the term Var(C) is computed in the same manner as Var(A) in Eq. (12).

Tolerance limits on the logarithm of the estimated failure rate are derived in the same manner as above with the last term of Eq. (10) replaced with

$$Z_{1-\beta/2}\left[(\Sigma\, \sigma_j^2 + \sigma_\epsilon^2)/\hat{\lambda}_s T_0\right]^{1/2} ,$$

where the term $\hat{\sigma}^2$ is the estimated error variance. The variance term in Eq. (10) changes since the estimated log failure rate is assumed to have a normal distribution with mean

$$\tau_s + \sum_{j\in I_1} B(j,s(j))$$

and variance

$$[\sum_{j \epsilon I_2} \sigma_j^2 + \sigma_\epsilon^2]/E(\lambda_s)T_o \quad .$$

The tolerance limits on the logarithm of the observed failure rate may be used to predict the failure rate of cells for which no data are available in the analysis; that is, for so-called future observations. We are at least $(1 - \alpha)\cdot 100\%$ confident that at least $(1 - \beta)\cdot 100\%$ of the logarithms of future failure rates for specific cells will be between the limits L and U for a specified time T_o.

Since tolerance limits are confidence limits on the percentiles of a statistical distribution, they do not shrink to zero as the sample size increases. However, as more and more sample data are included in the analysis, the uncertainty in the estimates decreases and hence, the terms Var(A) and Var(C) of Eqs. (11) and (12) also decrease to zero.

13. Plot of the Tolerance Interval for LN(FAILURE RATE). The tolerance intervals of the previous section are plotted. This gives the analyst some idea of cells in which there is a great degree of uncertainty in the logarithm of the failure rate on cells in which the failure rate is higher or lower than other cells. For example, the cell with index 10 in the simulated data seems to have a lower failure rate than other cells. This may be because cell 10 is the only cell with the horsepower at the lowest level.

14. Tolerance Intervals on Failure Rate by Cells (Tolerance Interval on Failure Rate Estimates by Cells). The tolerance intervals in this section are derived by exp(L) and exp(U), where L and U are the lower and upper tolerance interval limits discussed in Sec. III.B.12. The intervals in this section are tolerance intervals on the failure rate or the failure rate estimates for a time T_o. As such, these intervals reflect the uncertainty in the actual failure rates. In the example, we are at least 95% confident ($\alpha = 0.05$) that at least 90% ($\beta = 0.10$) of all of the true failure rates in 8760 hours for pumps under the same conditions as identified in the cell with index 1 will be between 5.832×10^{-5} and 4.484×10^{-3}.

15. Plot of Tolerance Interval for Failure Rate. The tolerance intervals of Sec. III.B.14 are plotted, analogous to the plotted intervals described in Sec. III.B.13.

16. Tolerance Intervals on LN(FAILURE RATE) for Unknown (Incompletely Specified) Cells. [Tolerance Intervals on LN(FAILURE RATE) Estimates for Unknown (Incompletely Specified) Cells]. Tolerance intervals are computed for prespecified cells (see Sec. III.A.4 for a description of the cell specification). The tolerance intervals of this section are derived exactly as those of Sec. III.B.12. Note that the intervals given in Sec. III.B.12 are for those cells for which there are data that are used in the analysis. The intervals given in this section need not be for cells for which there are no data.

17. Plot of Tolerance Intervals for LN(FAILURE RATE). Plots of the tolerance intervals of Sec. III.B.16 are given in this section.

18. "Average" Tolerance Intervals [LN(FAILURE RATE)]. Three types of average tolerance intervals are used to represent the "average" of the tolerance intervals of Sec. III.B.16. Let the m intervals of Sec. III.B.16 be denoted by (L_i,U_i), $i = 1,2, \ldots , m$, and the weights be w_i where $\Sigma w_i = 1$ (see Sec. III.A.4). Then the CONSERVATIVE (average) interval is given by the interval (L,U) where $L = \min L_i$ and $U = \max U_i$. The weighted ARITHMETIC average interval is given by $L = \Sigma w_i L_i$ and $U = \Sigma w_i U_i$. This ARITHMETIC AVERAGE interval is equivalent to a weighted geometric average of the corresponding tolerance intervals on the failure rate (see Sec. III.B.21). These two intervals are "ad hoc" intervals in the sense that they are not based on precise mathematical theory. The CONSERVATIVE interval by its construction should contain more than the specified proportion of the population having the desired (but unknown) characteristics. Such a conservative interval should be quite wide. The AVERAGE interval on the other hand will be shorter, but again no guarantee of a specified coverage can be provided.

Now let us consider an "average" interval in which an approximate guaranteed coverage is provided. Let the m cells for which intervals were desired for a time T_0 in Sec. III.B.16 be denoted by

$$S_i = \{(1,i(1)),(2,i(2)), \ldots , (k,i(k))\} ,$$

$i = 1,2, \ldots ,m$. Recall that for random factors $i(j) = 0$. Now, let us assume that both the log failure rate (in which $I_2 \neq 0$), as well as the logarithms of the failure rate estimates over all the specified cells, are mixtures of normal

distributions with weights w_i. (The weight w_i may be viewed as the prior probability that a failure rate will be in cell S_i.) Thus, the cumulative distribution function of the log failure rate or log failure rate estimate over the m cells is given by

$$\Sigma\, w_i\, \Phi\left(\frac{X - \mu_i}{\sigma_i^*}\right) \quad , \tag{14}$$

where $\Phi(Z)$ is the standard normal cumulative normal distribution function. The value of μ_i in Eq. (14) is given by

$$\mu_i = \tau_g + \sum_{j \epsilon I_1} B(j,i(j)) \quad , \quad i(j) \epsilon S_i \tag{15}$$

where τ_g, $B(j,i(j))$, and I_1 are defined in Sec. II. The variance in Eq. (14) is given by

$$\sigma_i^{*2} = \begin{cases} \left(\sum_{j \epsilon I_2} \sigma_j^2\right)/E(\lambda_i)T_o & \text{for log failure rates } (I_2 \neq 0) \\ \left(\sum_{j \epsilon I_2} \sigma_j^2 + \sigma_\varepsilon^2\right)/E(\lambda_i)T_o & \text{for log failure rate estimates} \end{cases} . \tag{16}$$

The quantities μ_i and σ_i^{*2} of Eqs. (15) and (16) are estimated by

$$\hat{\mu}_i = \hat{\tau}_g + \sum_{j \epsilon I_1} \hat{B}(j,i(j)) \quad , \quad i(j) \in S_i \tag{17}$$

and

$$\hat{\sigma}_i^{*2} = \begin{cases} \left(\sum_{j \epsilon I_1} \hat{\sigma}_j^2\right) \cdot e^{-\hat{\mu}_i}/T_o & \text{for log failure rates } (I_2 \neq 0) \\ \left(\sum_{j \epsilon I_2} \hat{\sigma}_j^2 + \hat{\sigma}_\varepsilon^2\right) \cdot e^{-\hat{\mu}_i}/T_o & \text{for log failure rate estimates} \end{cases} \tag{18}$$

where the estimates $\hat{B}$ and $\hat{\sigma}_j^{*2}$ are discussed in Sec. III.B.4 and 5, and $e^{-\hat{\mu}_i}$ serves as an estimate for $E(\lambda_i)$.

What we now desire is a $(1 - \alpha)100\%$, $(1 - \beta)100\%$ tolerance interval on the mixture populations. This interval is called the MIXTURE AVERAGE in the output. To obtain the desired interval, the two percentiles Z_L and Z_u are estimated by solving the pair of equations

$$\Sigma\, w_i\, \Phi\left(\frac{\hat{Z}_L - \hat{\mu}_i}{\hat{\sigma}_i^*}\right) = \beta/2 \quad \text{and} \quad \Sigma\, w_i\, \Phi\left(\frac{\hat{Z}_u - \hat{\mu}_i}{\hat{\sigma}_i^*}\right) = 1 - \beta/2 \tag{19}$$

for $\hat{Z}_L$ and $\hat{Z}_u$. Then the required tolerance interval on the mixture population is given by

$$\left(\hat{Z}_L - Z_{1-\alpha/2}[\mathrm{Var}(\hat{Z}_L)]^{1/2}\ ,\ \hat{Z}_u + Z_{1-\alpha/2}[\mathrm{Var}(\hat{Z}_u)]^{1/2}\right), \tag{20}$$

where $\mathrm{Var}(\hat{Z}_L)$ and $\mathrm{Var}(\hat{Z}_u)$ are determined by propagation of error as follows. (We consider here only $\hat{\mathrm{Var}}(Z_L)$, realizing $\hat{\mathrm{Var}}(Z_u)$ is found in a similar way.)

$$\begin{aligned}\mathrm{Var}(\hat{Z}_L) = {} & \sum_i \sum_j \mathrm{Cov}(\hat{\mu}_i, \hat{\mu}_j)(\partial \hat{Z}_L/\partial \hat{\mu}_i)(\partial \hat{Z}_L/\partial \hat{\mu}_j) \\ & + \sum_i \sum_j \mathrm{Cov}(\hat{\sigma}_i^{*2}, \hat{\sigma}_j^{*2})(\partial \hat{Z}_L/\partial \hat{\sigma}_i^{*2})(\partial \hat{Z}_L/\partial \hat{\sigma}_j^{*2}) \\ & + 2 \sum_i \sum_j \mathrm{Cov}(\mu_i, \hat{\sigma}_j^{*2})(\partial \hat{Z}_L/\partial \hat{\sigma}_j^{*2})(\partial \hat{Z}_L/\partial \hat{\mu}_i)\ . \end{aligned} \tag{21}$$

The partial derivatives in Eq. (21) are obtained by considering the function of $\hat{Z}_L$, $\hat{\mu}_i$, and $\hat{\sigma}_j^{*2}$, namely,

$$f = \Sigma\, w_i\, \Phi\left(\frac{\hat{Z}_L - \hat{\mu}_i}{\hat{\sigma}_i^*}\right) - \frac{\beta}{2} = \frac{1}{\sqrt{2\pi}} \Sigma\, w_i \int_{-\infty}^{\frac{\hat{Z}_L - \hat{\mu}_i}{\hat{\sigma}_i^*}} e^{-t^2} dt - \frac{\beta}{2} = 0\ . \tag{22}$$

Then by implicit differentiation of Eq. (22),

$$\partial\hat{Z}_L/\partial\hat{\mu}_i = -(\partial f/\partial\hat{\mu}_i)/(\partial f/\partial\hat{Z}_L)$$

and

$$\partial\hat{Z}_L/\partial\hat{\sigma}_i^{*2} = -(\partial f/\partial\hat{\sigma}_i^{*2})/(\partial f/\partial\hat{Z}_L) \quad .$$

The covariances in Eq. (21) are obtained by propagation of error using Eqs. (17) and (18), giving

$$\mathrm{Cov}(\hat{\mu}_i,\hat{\mu}_j) = \mathrm{Var}(\hat{\tau}_g) + \sum_{k\epsilon I_1}\sum_{\ell\epsilon I_1} \mathrm{Cov}(\hat{B}(k,i(k)), \hat{B}(\ell,j(\ell))$$

$$+ \sum_{k\epsilon I_1} \mathrm{Cov}(\hat{\tau}_g,\hat{B}(k,i(k))) + \sum_{\ell\epsilon I_1} \mathrm{Cov}(\hat{\tau}_g,\hat{B}(\ell,j(\ell)))$$

for $i(k) \epsilon S_i$ and $j(\ell) \epsilon S_j$. Letting

$$\hat{s}^2 = \sum_{j\epsilon I_2} \hat{\sigma}_j^2 \quad \text{or} \quad \sum_{j\epsilon I_2} \hat{\sigma}_j^2 + \hat{\sigma}_\epsilon^2 \ ,$$

we have that

$$\mathrm{Cov}(\hat{\sigma}_i^{*2},\hat{\sigma}_j^{*2}) = e^{-\hat{\mu}_i} \cdot e^{-\hat{\mu}_j}(\mathrm{Var}(\hat{s}^2) + (\hat{s}^2)^2\mathrm{Cov}(\hat{\mu}_i,\hat{\mu}_j))/T_o \ ,$$

and

$$\mathrm{Cov}(\hat{\mu}_i,\hat{\sigma}_j^{*2}) = -\hat{s}^2 e^{-\hat{\mu}_j} \mathrm{Cov}(\hat{\mu}_i,\hat{\mu}_j)/T_o \ .$$

The tolerance interval given by Eq. (20) and obtained by using the above equations is given in the listing as the MIXTURE AVERAGE interval. Since this interval has a precise statistical foundation with associated confidence and coverage, it is recommended in preference to the CONSERVATIVE or ARITHMETIC MEAN average intervals. In the example, the ARITHMETIC MEAN interval and MIXTURE AVERAGE interval are nearly equivalent, although this will not always be the case.

Los Alamos Identification No. 1392

①program title: frac (failure rate analysis code)

developed for: u.s. nuclear regulatory commission
office of nuclear regulatory research
division of risk analysis

date: november 1981

contacts: h. f. martz, group s-1, ms600, los alamos
national laboratory, los alamos, nm 87545

phone number: (505) 667-2687 (fts (505) 843-2687)

this program was prepared for an agency of
the united states government. neither the
united states government nor any agency thereof,
or any of their employees, makes any warranty,
expressed or implied, or assumes any legal
liability or responsibility for any third
party's use, or the results of such use, of
any portion of this program or represents
that its use by such third party would not
infringe privately owned rights.

```
/title  paragraph
    test routine with simulated data november 1981
/model paragraph
plant    random              1
sysw/plt    random           1,2
type  fixed  2               3
head  fixed  4                4
hspwr     fixed  4              5
typ*head fixed                  3,4
type*sys random                 1,2,3
/analysis paragraph
          0          1
/data paragraph
data is included in input file, format is 'free'
1,1,1,3,4  37  .14e+06
1,1,1,4,3  3  .175e+05
1,1,1,4,4  8  .175e+05
1,1,2,2,2  2   .175e+05
1,1,2,3,3  5   .175e+05
1,1,2,3,4 31   .105e+06
1,1,2,4,4 11   .350e+05
1,2,1,2,2  5   .175e+05
1,2,1,2,3 28  .701e+05
1,2,2,1,1  0   .175e+05
1,2,2,1,2  3   .175e+05
1,2,2,1,3 29   .526e+05
1,2,2,2,2 27   .701e+05
1,2,2,3,4 16   .175e+05
2,1,1,3,2  10   .175e+05
2,1,1,3,3 25  .175e+05
2,1,2,1,2  109   .526e05
2,1,2,2,2  43   .175e+05
2,1,2,2,3  1058  .298e+06
2,1,2,2,4  198   .350e+05
2,2,1,3,2  64   .701e+05
2,2,1,3,3 42  .350e+05
```

```
2,2,1,4,2   39  .35e+05
2,2,2,4,2   36  .35e+05
/tolerance paragraph
          1   5.0000000000000000e-02   0.1000000000000000
   8760.000000000000
0.0,1,3,3  .2
0.0,2,1,4  .4
0.0,1,2,4  .3
0.0,2,1,3  .1
/end
```

frac - failure rate analysis code

test routine with simulated data november 1981

② in failure rate analysis model...

	factor index	factor name	factor type	number of levels	factor subscripts
main factors	1	plant	random		1
	2	sysw/plt	random		1 2
	3	type	fixed	2	3
	4	head	fixed	4	4
	5	hspwr	fixed	4	5
interactive factors	6	typ*head	fixed	8	3 4
	7	type*sys	random		1 2 3

failure rate analysis

number of cases read.... 24

③ failure rate information by cells

index	cell	failures	time on test	failure rate	ln(f.r.)	weight
1	1, 1, 1, 3, 4, 3, 1	37	0.1400e+06	0.2643e-03	-.8238e+01	0.3700e+02
2	1, 1, 1, 4, 3, 4, 1	3	0.1750e+05	0.1714e-03	-.8671e+01	0.3000e+01
3	1, 1, 1, 4, 4, 4, 1	8	0.1750e+05	0.4571e-03	-.7691e+01	0.8000e+01
4	1, 1, 2, 2, 2, 6, 2	2	0.1750e+05	0.1143e-03	-.9077e+01	0.2000e+01
5	1, 1, 2, 3, 3, 7, 2	5	0.1750e+05	0.2857e-03	-.8161e+01	0.5000e+01
6	1, 1, 2, 3, 4, 7, 2	31	0.1050e+06	0.2952e-03	-.8128e+01	0.3100e+02
7	1, 1, 2, 4, 4, 8, 2	11	0.3500e+05	0.3143e-03	-.8065e+01	0.1100e+02
8	1, 2, 1, 2, 2, 2, 3	3	0.1750e+05	0.1714e-03	-.8671e+01	0.3000e+01
9	1, 2, 1, 2, 3, 2, 3	28	0.7010e+05	0.3994e-03	-.7825e+01	0.2800e+02
10	1, 2, 2, 1, 1, 5, 4	0	0.1750e+05	0.1905e-04	-.1087e+02	0.3333e+00
11	1, 2, 2, 1, 2, 5, 4	3	0.1750e+05	0.1714e-03	-.8671e+01	0.3000e+01
12	1, 2, 2, 1, 3, 5, 4	29	0.5260e+05	0.5513e-03	-.7503e+01	0.2900e+02
13	1, 2, 2, 2, 2, 6, 4	27	0.7010e+05	0.3852e-03	-.7862e+01	0.2700e+02
14	1, 2, 2, 3, 4, 7, 4	16	0.1750e+05	0.9143e-03	-.6997e+01	0.1600e+02
15	2, 1, 1, 3, 2, 3, 1	10	0.1750e+05	0.5714e-03	-.7467e+01	0.1000e+02
16	2, 1, 1, 3, 3, 3, 1	25	0.1750e+05	0.1429e-02	-.6551e+01	0.2500e+02
17	2, 1, 2, 1, 2, 5, 2	109	0.5260e+05	0.2072e-02	-.6179e+01	0.1090e+03

18	2. 1. 2. 2. 2. 8. 2	43	0.1750e+05	0.2457e-02	-.6009e+01	0.4300e+02
19	2. 1. 2. 2. 3. 8. 2	1058	0.2980e+06	0.3550e-02	-.5641e+01	0.1058e+04
20	2. 1. 2. 2. 4. 8. 2	198	0.3500e+05	0.5657e-02	-.5175e+01	0.1980e+03
21	2. 2. 1. 3. 2. 3. 3	64	0.7010e+05	0.9130e-03	-.6999e+01	0.6400e+02
22	2. 2. 1. 3. 3. 3. 3	42	0.3500e+05	0.1200e-02	-.6725e+01	0.4200e+02
23	2. 2. 1. 4. 2. 4. 3	39	0.3500e+05	0.1114e-02	-.6800e+01	0.3900e+02
24	2. 2. 2. 4. 2. 8. 4	36	0.3500e+05	0.1029e-02	-.6880e+01	0.3600e+02

④ estimates of ln of fixed effects

effect	estimate	std dev	est/(std dev)
constant	-.8040e+01	0.8095e+00	-.9933e+01
type	-.2463e+00	0.4874e-01	-.5052e+01
head	-.1590e-01	0.0000e+00	0.0000e+00
head	0.4552e+00	0.7292e-01	0.6243e+01
head	-.2186e+00	0.6681e-01	-.3272e+01
hspwr	-.2515e+01	0.1186e+01	-.2121e+01
hspwr	0.3718e+00	0.3977e+00	0.9348e+00
hspwr	0.8336e+00	0.3973e+00	0.2098e+01
typ*head	-.3504e+00	0.8244e-01	-.4251e+01
typ*head	0.2895e-01	0.7779e-01	0.3722e+00
typ*head	0.1806e-01	0.8382e-01	0.2155e+00

estimates of fixed effects

effect	estimate	std dev
constant	0.3222e-03	0.2608e-03
type	0.7817e+00	0.3810e-01
head	0.9842e+00	0.0000e+00
head	0.1577e+01	0.1150e+00
head	0.8037e+00	0.5369e-01
hspwr	0.8085e-01	0.9585e-01
hspwr	0.1450e+01	0.5769e+00
hspwr	0.2301e+01	0.9143e+00
typ*head	0.7044e+00	0.5807e-01
typ*head	0.1029e+01	0.8008e-01
typ*head	0.1018e+01	0.8534e-01

⑤ random effect variance estimates

effect	estimate	std dev	est/(std dev)
err.var.	0.8281e+00	0.2556e+00	0.3240e+01
plant	0.9762e+00	0.9873e+00	0.9888e+00
sysw/plt	0.1714e-01	0.2702e-01	0.6346e+00
type*sys	0.0000e+00	0.0000e+00	0.0000e+00

⑥ -2*log(maximum likelihood) = 0.7892e+02

⑦ variance-covariance matrix for fixed effects

	constant	type	head	head	head	hspwr	hspwr	hspwr	typ*head
constant	0.6552e+00								
type	0.1743e-02	0.2376e-02							
head	0.0000e+00	0.0000e+00	0.0000e+00						
head	0.1203e-03	0.1450e-02	0.0000e+00	0.5318e-02					
head	-.5244e-03	-.1208e-02	0.0000e+00	-.1689e-02	0.4464e-02				
hspwr	0.4686e+00	0.2370e-02	0.0000e+00	0.4187e-03	0.5479e-03	0.1406e+01			
hspwr	-.1556e+00	-.3868e-03	0.0000e+00	0.5276e-03	0.2405e-03	-.4670e+00	0.1582e+00		
hspwr	-.1564e+00	-.1028e-02	0.0000e+00	-.1094e-02	-.2436e-04	-.4691e+00	0.1549e+00	0.1578e+00	
typ*head	0.3164e-02	0.8183e-03	0.0000e+00	.1057e-02	0.9971e-03	0.5181e-02	0.1015e-03	-.2100e-02	0.6797e-02
typ*head	0.1587e-02	0.7677e-03	0.0000e+00	0.3098e-02	-.7986e-03	0.1102e-02	-.1029e-02	-.2987e-03	-.3094e-03
typ*head	-.2917e-02	-.1331e-02	0.0000e+00	-.7066e-03	-.1454e-02	-.3878e-02	0.5755e-03	0.1405e-02	-.3422e-02

	typ*head	typ*head
typ*head	0.6051e-02	
typ*head	-.2863e-02	0.7025e-02

⑧
variance-covariance matrix for random effects

	err.var.	plant	sysw/plt	type*sys
err.var.	0.6531e-01			
plant	-.1232e-03	0.9748e+00		
sysw/plt	-.1541e-03	-.3593e-03	0.7299e-03	
type*sys	0.0000e+00	0.0000e+00	0.0000e+00	0.0000e+00

⑨
design matrix (x)

index	dummy variables										
1	1.	1.	0.	0.	1.	-1.	-1.	-1.	0.	0.	1.
2	1.	1.	-1.	-1.	-1.	0.	0.	1.	-1.	-1.	-1.
3	1.	1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
4	1.	-1.	0.	1.	0.	0.	1.	0.	0.	-1.	0.
5	1.	-1.	0.	0.	1.	0.	0.	1.	0.	0.	-1.
6	1.	-1.	0.	0.	1.	-1.	-1.	-1.	0.	0.	-1.
7	1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	1.	1.	1.
8	1.	1.	0.	1.	0.	0.	1.	0.	0.	1.	0.
9	1.	1.	0.	1.	0.	0.	0.	1.	0.	1.	0.
10	1.	-1.	1.	0.	0.	1.	0.	0.	-1.	0.	0.
11	1.	-1.	1.	0.	0.	0.	1.	0.	-1.	0.	0.
12	1.	-1.	1.	0.	0.	0.	0.	1.	-1.	0.	0.
13	1.	-1.	0.	1.	0.	0.	1.	0.	0.	-1.	0.
14	1.	-1.	0.	0.	1.	-1.	-1.	-1.	0.	0.	-1.
15	1.	1.	0.	0.	1.	0.	1.	0.	0.	0.	1.
16	1.	1.	0.	0.	1.	0.	0.	1.	0.	0.	1.
17	1.	-1.	1.	0.	0.	0.	1.	0.	-1.	0.	0.
18	1.	-1.	0.	1.	0.	0.	1.	0.	0.	-1.	0.
19	1.	-1.	0.	1.	0.	0.	0.	1.	0.	-1.	0.
20	1.	-1.	0.	1.	0.	-1.	-1.	-1.	0.	-1.	0.
21	1.	1.	0.	0.	1.	0.	1.	0.	0.	0.	1.
22	1.	1.	0.	0.	1.	0.	0.	1.	0.	0.	1.
23	1.	1.	-1.	-1.	-1.	0.	1.	0.	-1.	-1.	-1.
24	1.	-1.	-1.	-1.	-1.	0.	1.	0.	1.	1.	1.

⑩
ln(failure rate) residual analysis

index	observed ln failure rate	predicted ln failure rate	std dev predicted	95% confidence interval	residual (o-p)	std dev residual	standardized residual
1	-.8238e+01	-.7177e+01	0.7101e+00	(-.8569e+01,-.5785e+01)	-.1061e+01	0.1355e+01	-.7835e+00
2	-.8671e+01	-.7370e+01	0.7195e+00	(-.8780e+01,-.5960e+01)	-.1301e+01	0.1412e+01	-.9214e+00
3	-.7691e+01	-.6894e+01	0.7198e+00	(-.8305e+01,-.5483e+01)	-.7966e+00	0.1373e+01	-.5800e+00
4	-.9077e+01	-.6996e+01	0.7112e+00	(-.8390e+01,-.5602e+01)	-.2081e+01	0.1440e+01	-.1445e+01

5	-.8161e+01	-.7197e+01	0.7192e+00	(-.8607e+01,-.5787e+01)	-.9635e+00	0.1387e+01	-.6945e+00
6	-.8128e+01	-.6721e+01	0.7182e+00	(-.8128e+01,-.5313e+01)	-.1407e+01	0.1356e+01	-.1038e+01
7	-.8065e+01	-.7008e+01	0.7202e+00	(-.8420e+01,-.5597e+01)	-.1057e+01	0.1367e+01	-.7732e+00
8	-.8671e+01	-.7430e+01	0.7292e+00	(-.8860e+01,-.6001e+01)	-.1241e+01	0.1414e+01	-.8777e+00
9	-.7825e+01	-.6969e+01	0.7257e+00	(-.8391e+01,-.5546e+01)	-.8568e+00	0.1357e+01	-.6316e+00
10	-.1087e+02	-.9975e+01	0.1727e+01	(-.1336e+02,-.6589e+01)	-.8939e+00	0.3282e+01	-.2723e+00
11	-.8671e+01	-.7088e+01	0.7098e+00	(-.8479e+01,-.5696e+01)	-.1584e+01	0.1410e+01	-.1123e+01
12	-.7503e+01	-.6626e+01	0.7124e+00	(-.8022e+01,-.5230e+01)	-.8773e+00	0.1356e+01	-.6470e+00
13	-.7862e+01	-.6996e+01	0.7112e+00	(-.8390e+01,-.5602e+01)	-.8660e+00	0.1357e+01	-.6384e+00
14	-.6997e+01	-.6721e+01	0.7182e+00	(-.8128e+01,-.5313e+01)	-.2766e+00	0.1362e+01	-.2032e+00
15	-.7467e+01	-.8115e+01	0.7110e+00	(-.9509e+01,-.6722e+01)	0.6478e+00	0.1368e+01	0.4735e+00
16	-.6551e+01	-.7653e+01	0.7095e+00	(-.9044e+01,-.6263e+01)	0.1102e+01	0.1357e+01	0.8123e+00
17	-.6179e+01	-.7088e+01	0.7098e+00	(-.8479e+01,-.5696e+01)	0.9085e+00	0.1351e+01	0.6723e+00
18	-.6009e+01	-.6996e+01	0.7112e+00	(-.8390e+01,-.5602e+01)	0.9871e+00	0.1354e+01	0.7291e+00
19	-.5641e+01	-.6534e+01	0.7074e+00	(-.7921e+01,-.5148e+01)	0.8934e+00	0.1350e+01	0.6619e+00
20	-.5175e+01	-.6058e+01	0.7092e+00	(-.7448e+01,-.4668e+01)	0.8830e+00	0.1351e+01	0.6538e+00
21	-.6999e+01	-.8115e+01	0.7110e+00	(-.9509e+01,-.6722e+01)	0.1116e+01	0.1353e+01	0.8254e+00
22	-.6725e+01	-.7653e+01	0.7095e+00	(-.9044e+01,-.6263e+01)	0.9280e+00	0.1354e+01	0.6853e+00
23	-.6800e+01	-.7832e+01	0.7182e+00	(-.9240e+01,-.6424e+01)	0.1032e+01	0.1355e+01	0.7622e+00
24	-.6880e+01	-.7946e+01	0.7190e+00	(-.9356e+01,-.6537e+01)	0.1087e+01	0.1355e+01	0.7873e+00

(11)

failure rate residual analysis

index	observed failure rate	predicted failure rate	95% confidence interval	residual (o-p)
1	0.2643e-03	0.7639e-03	(0.1899e-03,0.3073e-02)	-.4996e-03
2	0.1714e-03	0.6297e-03	(0.1537e-03,0.2580e-02)	-.4583e-03
3	0.4571e-03	0.1014e-02	(0.2473e-03,0.4157e-02)	-.5568e-03
4	0.1143e-03	0.9156e-03	(0.2271e-03,0.3691e-02)	-.8014e-03
5	0.2857e-03	0.7488e-03	(0.1829e-03,0.3066e-02)	-.4631e-03
6	0.2952e-03	0.1206e-02	(0.2950e-03,0.4927e-02)	-.9104e-03
7	0.3143e-03	0.9044e-03	(0.2204e-03,0.3710e-02)	-.5901e-03
8	0.1714e-03	0.5929e-03	(0.1420e-03,0.2476e-02)	-.4215e-03
9	0.3994e-03	0.9409e-03	(0.2269e-03,0.3902e-02)	-.5414e-03
10	0.1905e-04	0.4657e-04	(0.1576e-05,0.1376e-02)	-.2752e-04
11	0.1714e-03	0.8354e-03	(0.2078e-03,0.3358e-02)	-.6640e-03
12	0.5513e-03	0.1326e-02	(0.3281e-03,0.5356e-02)	-.7743e-03
13	0.3852e-03	0.9156e-03	(0.2271e-03,0.3691e-02)	-.5305e-03
14	0.9143e-03	0.1206e-02	(0.2950e-03,0.4927e-02)	-.2914e-03
15	0.5714e-03	0.2990e-03	(0.7420e-04,0.1205e-02)	0.2725e-03
16	0.1429e-02	0.4744e-03	(0.1181e-03,0.1906e-02)	0.9542e-03
17	0.2072e-02	0.8354e-03	(0.2078e-03,0.3358e-02)	0.1237e-02
18	0.2457e-02	0.9156e-03	(0.2271e-03,0.3691e-02)	0.1541e-02
19	0.3550e-02	0.1453e-02	(0.3632e-03,0.5814e-02)	0.2097e-02
20	0.5657e-02	0.2340e-02	(0.5827e-03,0.9393e-02)	0.3318e-02
21	0.9130e-03	0.2990e-03	(0.7420e-04,0.1205e-02)	0.6140e-03
22	0.1200e-02	0.4744e-03	(0.1181e-03,0.1906e-02)	0.7256e-03
23	0.1114e-02	0.3968e-03	(0.9711e-04,0.1622e-02)	0.7175e-03
24	0.1029e-02	0.3540e-03	(0.8648e-04,0.1449e-02)	0.6746e-03

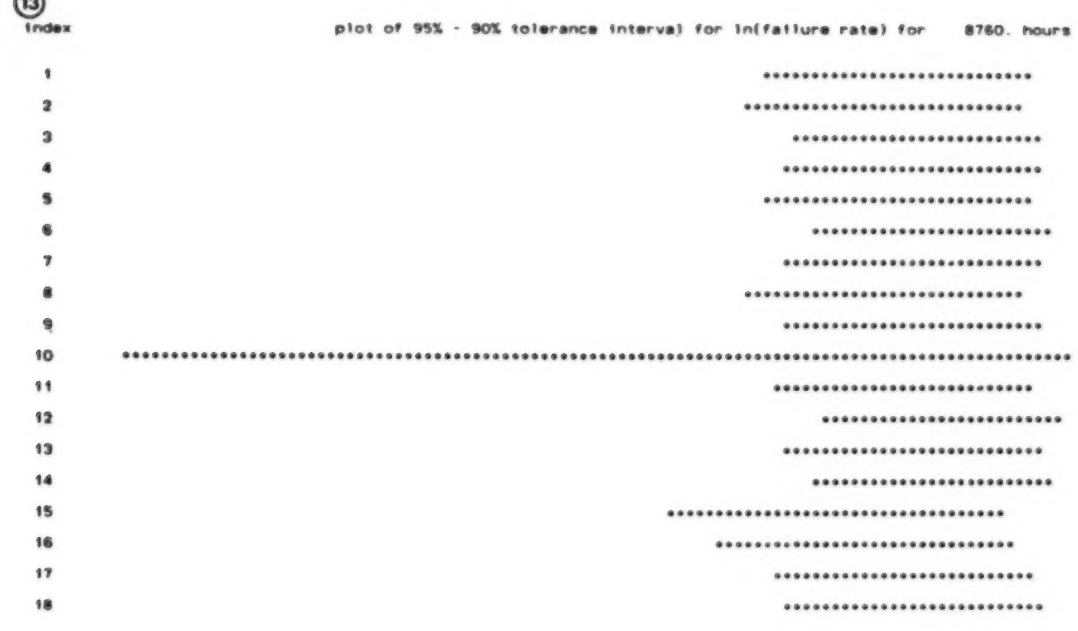
(12)

tolerance intervals on ln(failure rate) by cells for 8760.hours

index	cell	95% - 90% tolerance interval
1	1, 1, 1, 3, 4, 3, 1	(-.9749e+01,-.5407e+01)
2	1, 1, 1, 4, 3, 4, 1	(-.1009e+02,-.5527e+01)

3	1, 1, 1, 4, 4, 4, 1	(-.9325e+01,-.5187e+01)
4	1, 1, 2, 2, 2, 6, 2	(-.9463e+01,-.5275e+01)
5	1, 1, 2, 3, 3, 7, 2	(-.9804e+01,-.5411e+01)
6	1, 1, 2, 3, 4, 7, 2	(-.9059e+01,-.5052e+01)
7	1, 1, 2, 4, 4, 8, 2	(-.9504e+01,-.5273e+01)
8	1, 2, 1, 2, 2, 2, 3	(-.1022e+02,-.5556e+01)
9	1, 2, 1, 2, 3, 2, 3	(-.9455e+01,-.5237e+01)
10	1, 2, 2, 1, 1, 5, 4	(-.2069e+02,-.4715e+01)
11	1, 2, 2, 1, 2, 5, 4	(-.9604e+01,-.5344e+01)
12	1, 2, 2, 1, 3, 5, 4	(-.8905e+01,-.4983e+01)
13	1, 2, 2, 2, 2, 6, 4	(-.9463e+01,-.5275e+01)
14	1, 2, 2, 3, 4, 7, 4	(-.9059e+01,-.5052e+01)
15	2, 1, 1, 3, 2, 3, 1	(-.1145e+02,-.5895e+01)
16	2, 1, 1, 3, 3, 3, 1	(-.1055e+02,-.5704e+01)
17	2, 1, 2, 1, 2, 5, 2	(-.9604e+01,-.5344e+01)
18	2, 1, 2, 2, 2, 6, 2	(-.9463e+01,-.5275e+01)
19	2, 1, 2, 2, 3, 6, 2	(-.8760e+01,-.4915e+01)
20	2, 1, 2, 2, 4, 6, 2	(-.8102e+01,-.4502e+01)
21	2, 2, 1, 3, 2, 3, 3	(-.1145e+02,-.5895e+01)
22	2, 2, 1, 3, 3, 3, 3	(-.1056e+02,-.5704e+01)
23	2, 2, 1, 4, 2, 4, 3	(-.1092e+02,-.5785e+01)
24	2, 2, 2, 4, 2, 8, 4	(-.1114e+02,-.5833e+01)

(13)



(14)

tolerance interval on failure rate by cells for 8760. hours

index	cell	95% - 90% tolerance interval
1	1, 1, 1, 3, 4, 3, 1	(0.5832e-04,0.4484e-02)
2	1, 1, 1, 4, 3, 4, 1	(0.4136e-04,0.3977e-02)
3	1, 1, 1, 4, 4, 4, 1	(0.8915e-04,0.5589e-02)
4	1, 1, 2, 2, 2, 6, 2	(0.7768e-04,0.5118e-02)
5	1, 1, 2, 3, 3, 7, 2	(0.3522e-04,0.4467e-02)
6	1, 1, 2, 3, 4, 7, 2	(0.1163e-03,0.6394e-02)
7	1, 1, 2, 4, 4, 8, 2	(0.7455e-04,0.5127e-02)
8	1, 2, 1, 2, 2, 2, 3	(0.3641e-04,0.3865e-02)
9	1, 2, 1, 2, 3, 2, 3	(0.7827e-04,0.5317e-02)
10	1, 2, 2, 1, 1, 5, 4	(0.1034e-08,0.8956e-02)
11	1, 2, 2, 1, 2, 5, 4	(0.6743e-04,0.4776e-02)
12	1, 2, 2, 1, 3, 5, 4	(0.1357e-03,0.6851e-02)
13	1, 2, 2, 2, 2, 6, 4	(0.7768e-04,0.5118e-02)
14	1, 2, 2, 3, 4, 7, 4	(0.1163e-03,0.6394e-02)
15	2, 1, 1, 3, 2, 3, 1	(0.1060e-04,0.2752e-02)
16	2, 1, 1, 3, 3, 3, 1	(0.2582e-04,0.3333e-02)
17	2, 1, 2, 1, 2, 5, 2	(0.6743e-04,0.4776e-02)
18	2, 1, 2, 2, 2, 6, 2	(0.7768e-04,0.5118e-02)
19	2, 1, 2, 2, 3, 6, 2	(0.1569e-03,0.7339e-02)
20	2, 1, 2, 2, 4, 6, 2	(0.3028e-03,0.1109e-01)
21	2, 2, 1, 3, 2, 3, 3	(0.1060e-04,0.2752e-02)
22	2, 2, 1, 3, 3, 3, 3	(0.2582e-04,0.3333e-02)
23	2, 2, 1, 4, 2, 4, 3	(0.1814e-04,0.3073e-02)
24	2, 2, 2, 4, 2, 8, 4	(0.1453e-04,0.2930e-02)

(15)

index plot of 95% - 90% tolerance interval for failure rate for 8760. hours

```
 8      ************************************
 9      ****************************************************
10      ************************************************************************************
11      ************************************************
12       ************************************************************
13      ************************************************
14       ************************************************************
15      **************************
16      *******************************
17      *********************************************
18      ************************************************
19       ********************************************************************
20        ************************************************************************************************************
21      **************************
22      *******************************
23      *******************************
24      ****************************
```

(16)

tolerance intervals on ln(failure rate) for unknown (incompletely specified) cells for 8760. hours

index	cell	95% - 90% tolerance interval	weight
1	0, 0, 1, 3, 3, 3, 0	(-.1056e+02,-.5704e+01)	.200
2	0, 0, 2, 1, 4, 5, 0	(-.8239e+01,-.4575e+01)	.400
3	0, 0, 1, 2, 4, 2, 0	(-.8751e+01,-.4849e+01)	.300
4	0, 0, 2, 1, 3, 5, 0	(-.8905e+01,-.4983e+01)	.100

(17)

```
index                          plot of 95% - 90% tolerance interval for ln(failure rate) for    8760. hours

  1                    ********************************************************************************
  2                                              ************************************************************
  3                                           ***************************************************************
  4                                         ****************************************************************
```

"average" tolerance intervals for 8760. hours

(18)

conservative (min - max)

arithmetic mean

mixture average

"average" values: conservative (-.1056e+02,-.4575e+01)
arithmetic mean (-.8924e+01,-.4924e+01)
mixture average (-.9945e+01,-.4464e+01)

(19) tolerance intervals on failure rate for unknown (incompletely specified) cells for 8760. hours

index	cell	95% - 90% tolerance interval	weight
1	0, 0, 1, 3, 3, 3, 0	(0.2582e-04,0.3333e-02)	.200
2	0, 0, 2, 1, 4, 5, 0	(0.2642e-03,0.1031e-01)	.400
3	0, 0, 1, 2, 4, 2, 0	(0.1583e-03,0.7834e-02)	.300
4	0, 0, 2, 1, 3, 5, 0	(0.1357e-03,0.6851e-02)	.100

(20) index plot of 95% - 90% tolerance interval for failure rate for 8760. hours

1

2

3

4

(21) "average" tolerance intervals for 8760. hours

conservative (min - max)

geometric mean

mixture average

"average" values: conservative (0.2582e-04,0.1031e-01)
arithmetic mean (0.1331e-03,0.7271e-02)
mixture average (0.4799e-04,0.1151e-01)

19. Tolerance Intervals on Failure Rate for Unknown (Incompletely Specified) Cells [Tolerance Intervals on Failure Rate Estimates for Unknown (Incompletely Specified) Cells]. Tolerance intervals given in this section are tolerance intervals on the failure rates rather than the log-failure rates. They are obtained by [exp(L),exp(U)], where L and U are the upper and lower tolerance bounds of Sec. III.B.16.

20. Plot of Tolerance Intervals for Failure Rate. The tolerance intervals of Sec. III.B.19 are plotted.

21. "Average" Tolerance Intervals (Failure Rate). From the three types of average intervals on the log-failure rates given in Sec. III.B.18, average tolerance intervals are derived for failure rates on failure rate estimates. If (L,U) represents an interval given in Sec. III.B.18, then the corresponding interval here is [exp(L),exp(U)].

One thing should be noted. The exponentiation of the weighted ARITHMETIC MEAN interval of the log-failure rates is equivalent to finding the weighted geometric mean of the failure rates, and for this reason the term GEOMETRIC MEAN is used in the output.

REFERENCES

1. H. F. Martz and R. J. Beckman, "Nuclear Reactor Component Populations, Reliability Data Bases, and Their Relationship to Failure Rate Estimation and Uncertainty Analysis," Nuclear Regulatory Commission report NUREG/CR-2433 (Los Alamos National Laboratory report LA-9115-MS) (1981).

2. A. H. Bowker and G. J. Lieberman, Engineering Statistics, 2nd Edition (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1972).

3. C. R. Hicks, Fundamental Concepts in the Design of Experiments (Holt, Rinehart, and Winston, New York, NY, 1973).

4. H. Scheffe, Analysis of Variance (John Wiley and Sons, Inc., New York, NY, 1959).

5. H. F. Martz, R. J. Beckman, and C. R. McInteer, "A Programmer's Guide to FRAC (Failure Rate Analysis Code)," Los Alamos National Laboratory internal report S-1/82-154 (1982).

6. H. O. Hartley and J. N. K. Rao, "Maximum-Likelihood Estimation for the Mixed Analysis of Variance Model," Biometrika 54, 93-108 (1967).

7. R. I. Jennrich and P. F. Sampson, "Newton-Raphson and Related Algorithms for Maximum Likelihood Variance3 Component Estimation," Technometrics 18, 11-17 (1976).

8. D. A. Harville, "Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems," J. Am. Stat. Assoc. 72, 320-338 (1977).

9. E. L. Welker and M. Lipow, "Estimating the Exponential Failure Rate from Data with No Failure Events," Proc. 1974 Annual Reliability and Maintainability Symposium, IEEE 74 CH0820-IRQC (1974).

10. W. J. Dixon and M. B. Brown, Editors, BMDP Biomedical Computer Programs, P-Series (University of California Press, Berkeley, California, 1979).

DISTRIBUTION

	Copies
Nuclear Regulatory Commission, AN, XA, RG, Bethesda, Maryland	323
Technical Information Center, Oak Ridge, Tennessee	2
Los Alamos National Laboratory, Los Alamos, New Mexico	50
	375